BMJ Open

Symptoms of COVID-19 in a population-based cohort study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-053403
Article Type:	Original research
Date Submitted by the Author:	17-May-2021
Complete List of Authors:	Khan, Sana; The University of Arizona, Department of Epidemiology and Biostatistics Farland, Leslie V.; The University of Arizona, Department of Epidemiology and Biostatistics; The University of Arizona Cancer Center Austhof, Erika; The University of Arizona, Department of Epidemiology and Biostatistics Bell, Melanie L.; The University of Arizona, Department of Epidemiology and Biostatistics Catalfamo, Collin; The University of Arizona, Department of Epidemiology and Biostatistics Chen, Zhao; The University of Arizona, Department of Epidemiology and Biostatistics Cordova-Marks, Felina; The University of Arizona Cancer Center; The University of Arizona, Department of Health Promotion Sciences Ernst, Kacey; The University of Arizona, Department of Epidemiology and Biostatistics, College of Public Health Garcia-Filion, Pamela; The University of Arizona, Department of Biomedical Informatics Heslin, Kelly M.; The University of Arizona, Department of Epidemiology and Biostatistics Hoskinson, Joshua; The University of Arizona Cancer Center Jehn, Megan; Arizona State University Joseph, Emily C.S.; The University of Arizona Cancer Center Kelley, Connor; The University of Arizona, Department of Epidemiology and Biostatistics Klimentidis, Yann; The University of Arizona, Department of Epidemiology and Biostatistics Russo Carroll, Stephanie; The University of Arizona, Department of Community, Environment, and Policy; The University of Arizona, Native Nations Institute at the Udall Center for Studies in Public Policy Kohler, Lindsay; The University of Arizona, Department of Epidemiology and Biostatistics; The University of Arizona, Department of Health Promotion Sciences Pogreba-Brown, Kristen; The University of Arizona, Department of Epidemiology and Biostatistics Jacobs, Elizabeth; The University of Arizona, Department of Epidemiology and Biostatistics; The University of Arizona Cancer Center
Keywords:	COVID-19, INFECTIOUS DISEASES, EPIDEMIOLOGY

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Symptoms of COVID-19 in a population-based cohort study

Sana M. Khan[1]

Leslie V. Farland[1,2]

Erika Austhof[1]

Melanie L. Bell[1]

Collin J. Catalfamo[1]

Zhao Chen[1]

Felina Cordova-Marks[2,3]

Kacey C. Ernst[1]

Pamela Garcia-Filion[4]

Kelly M. Heslin[1]

Joshua Hoskinson[2]

Megan L Jehn[5]

Emily C.S. Joseph[2]

Connor P. Kelley[1]

Yann Klimentidis[1]

Stephanie Russo Carroll[6,7]

Lindsay N. Kohler[1,3]

Kristen Pogreba-Brown[1]

Elizabeth T. Jacobs[1,2]

[1]Department of Epidemiology and Biostatistics, Mel and Enid Zuckerman College of Public Health, The University of Arizona, Tucson, AZ, United States; [2]University of Arizona Cancer Center, Tucson, AZ, United States, [3]Department of Health Promotion Sciences, Mel and Enid Zuckerman College of Public Health, The University of Arizona, Tucson, AZ, United States; [4]Department of Biomedical Informatics, College of Medicine–Phoenix, The University of Arizona, Phoenix, AZ, United States, [5]School of Human Evolution and Social Change, Arizona State University, Tempe, AZ, United States; [6]Department of Community, Environment and Policy, Mel and Enid Zuckerman College of Public Health, The University of Arizona, Tucson, AZ, United States; [7] Native Nations Institute at the Udall Center for Studies in Public Policy, University of Arizona, Tucson, Arizona.

Corresponding Author

Sana M. Khan, MPH
Doctoral Student, Epidemiology
University of Arizona
Mel and Enid Zuckerman College of Public Health
sanakhan@email.arizona.edu

Keywords: SARS-CoV-2, COVID-19, symptoms, prospective cohort

Abstract

Objective: To elucidate the symptom profile of laboratory-confirmed COVID-19 cases (16.2%) as compared to laboratory-confirmed negative individuals (22.4%) and to the untested general population (61.4%) to aid in earlier identification of SARS-CoV-2 infection. **Setting and Design:** We conducted a prospective, population-based cohort study, named The Arizona CoVHORT, among Arizona residents. **Participants:** Among the 1514 study participants, those who were COVID-19 positive were more likely to be Hispanic (33.5%) compared to COVID-19 negative participants (19.2%) and untested CoVHORT participants (13.8%), as well as more likely to report a body mass index (BMI) of $\geq$ 30 kg/m^2 (34.7%) compared with COVID-19 negative participants (31.0%) and untested CoVHORT participants (23.8%). In addition, those with COVID-19 were more likely to be never-smokers (93.5 vs. 86.1 and 90.3 for negative and untested participants, respectively). **Primary and Secondary Outcome Measures:** Of the 245 laboratory-confirmed COVID-19 cases, 38 (15.0%) reported having had no symptoms. Of those that did report symptoms, the most commonly reported ones were sore throat (19.0%), headache (15.5%), cough (12.7%), runny nose/cold-like symptoms (12.1%), and fatigue (12.0%). **Results:** In adjusted logistic regression models, COVID-19 positive participants were more likely than negative participants to experience loss of taste and smell (OR 35.7; 95% CI 18.4-69.5); bone or nerve pain (OR 17.9; 95% CI 6.7-47.4), vomiting (OR 10.8; 95% CI 3.1-37.5), nausea (OR 10.5; 95% CI 5.5-19.9), and headache (OR 8.4; 95% CI 5.6-12.8). **Conclusion:** When comparing confirmed COVID-19 cases with either confirmed negative or untested participants, the pattern of symptoms that discriminates SARS-CoV-2 infection from those arising from other potential circulating pathogens may differ from general reports of symptoms among cases alone.

Strengths and limitations of this study:

- The findings of this study will aid in the identification of symptoms that differentiate COVID-19 from other circulating infections or conditions, which is useful for resource-limited settings where diagnostic testing is limited.
- To our knowledge, no prior research has compared the prevalence of non-specific symptoms such as headache, fever, and runny nose between confirmed COVID-19-positive cases, confirmed COVID-19-negative cases, and population-based comparison groups.
- We cannot know the COVID-19 status of the untested participants; it is possible that some had already been infected but were asymptomatic or exhibited few symptoms.

Introduction

In late 2019, the novel coronavirus SARS-CoV-2 was first recognized in China among patients who presented with pneumonia and the first scientific report appeared shortly thereafter (1). On March 11th, 2020 the World Health Organization declared COVID-19 a pandemic. The pathogen has had multiple impacts on individual and societal wellbeing arising from both biological effects of the virus and policy-based mitigation. The majority of those infected with acute COVID-19 will go on to recover, though approximately 10-20% of COVID-19 patients overall will develop a severe case of disease, and may suffer from stroke, pneumonia, or acute respiratory distress syndrome (ARDS) and require intensive care and ventilation (2, 3).

Individuals are likely be most infectious during the early phases of the disease, when symptoms may be comparatively mild; therefore, it is important to elucidate the reported symptom patterns of COVID-19 patients compared to both laboratory-confirmed negative individuals and population-based controls. Several risk factors have been associated with disease susceptibility and severity including increasing age (4), male sex (2, 5, 6), and current or former smoking (3), which may also affect symptomology. Further, important differences in disease incidence and severity by race and ethnicity have emerged, with Native Americans, African-Americans, and Latinos having higher COVID-19 prevalence, hospitalization, and mortality rates compared to non-Hispanic whites (7). It is presently not known if reports of symptoms or symptom patterns vary by these factors as well.

A recent meta-analysis of over 24,000 patients across nine countries reported on COVID-19 symptom presentation. In this work, the most commonly-reported symptoms among people with COVID-19 were fever (78% of COVID-19 patients reporting), cough (57%), and fatigue (31%) (8). In comparison, another study conducted among European patients (n=1420) with mild or moderate COVID-19 found that the most frequently reported symptoms were headache (70%), loss of smell (70%), and obstruction of the nasal passages (68%) (9). The authors of another study, the objective of which was to develop a better symptom modeling algorithm to aid targeted testing, concluded that fever and cough should be used as the key symptoms for rapid COVID-19 screening given their high sensitivity (10). However, a major limitation of studies conducted to date is the lack of comparison of patient-reported symptoms to those of uninfected individuals. To our knowledge, no prior research has compared the prevalence of non-specific symptoms such as headache, fever, and runny nose between confirmed COVID-19-positive cases, confirmed COVID-19-negative cases, and population-based comparison groups.

Since COVID-19 community transmission began, Arizona has twice experienced severe COVID-19 surges, with more than 850,000 infections and 16,000 COVID-19-related deaths as of March, 2021. To address this epidemiological challenge, in May 2020, we initiated a large, prospective, population-based cohort in Arizona of racially- and ethnically- diverse residents in order to rigorously investigate factors contributing to variability in natural COVID-19 disease history including incidence, progression, resolution, and chronic outcomes of infection (11). This COVID-19 cohort, dubbed The Arizona CoVHORT, provides a rich data source for multiple areas of inquiry related to the pandemic. The objective of the present work was to determine which symptoms were reported with the greatest frequency among participants who tested positive for

COVID-19 as compared to participants who tested negative for COVID-19 and untested participants, while controlling for potential confounders such as age, ethnicity, and sex. The findings of this paper will aid in the identification of symptoms that differentiate COVID-19 from other circulating infections or conditions, such as allergies.

Materials and Methods

Study Participants

The overall goal of the CoVHORT is to continuously enroll Arizonans into a cohort study to track both the acute and long-term phases of infection with SARS-CoV-2. The present analysis includes data through October 31st, 2020, five months since the cohort was launched on May 28th, 2020. Several recruitment methods were employed, which have been described in detail previously (11). Briefly, the primary sources of recruitment have been through case investigations in a partnership with the Arizona Department of Health Services and the COVID-19 Antibody Testing Initiative (CATI), both of which have allowed for inclusion of laboratory-confirmed COVID-19 positive participants. By October 31st, 2020, a total of 176 COVID-19-positive participants had been recruited through health department case investigations and 10 through our partnership with CATI; further, a total of 168 participants who had COVID-19-negative results were recruited via CATI.

A comprehensive mailing list was purchased that provides information on 2.2 million residents in Arizona. To recruit the population-based comparison group, a total of 17,500 postcards were mailed to a simple random sample of Pima County, Arizona residents in July 2020. Consistent with the Dillman method to maximize participation and minimize bias (12), three phased mailings of recruitment postcards occurred every two weeks. Participant-provided information from baseline surveys was used to exclude those who had already enrolled from subsequent phases of the mailing campaign. Each list was screened prior to each mailing to reduce the number of undeliverable postcards. We have completed all three phases of the mailing campaign in Pima County, with 17,294 postcards delivered in the first phase, 17,147 in the second phase, and another 17,081 in the third phase. Method of recruitment is recorded for all participants allowing sensitivity analyses to be conducted within subgroups.

Patient and Public Involvement

We encourage active participation from members of The Arizona CoVHORT. The public and members of the cohort are invited to webinars where they are able to provide input, ask questions, and speak with the projects' principal investigators. We regularly revisit our survey instruments to ensure they are reflecting feedback from participants and are centering their experiences and priorities. Study findings will be disseminated at our study website (covhort.arizona.edu), along with a regularly updated participant dashboard containing descriptive data of the cohort population.

Survey Instruments

All participants included in the CoVHORT were sent identical structured electronic questionnaires at baseline, regardless of COVID-19 status. Participants were asked if they had tested for the virus that causes COVID-19 with a nasal swab, throat swab, or saliva test since January 2020.

Participants were classified as untested, positive or negative based on their results. Participants, regardless of COVID-19 test status, were asked, "Since January 1, have you experienced a sudden illness that led you to believe you had COVID19?" If they answered "yes", all participants, regardless of case status, were asked to indicate which symptoms they had experienced since January 2020 from a list based upon prior reports in the literature, as well as through an open-text field. Further, COVID-19 positive participants were queried regarding the first symptom that they recalled having experienced. Information regarding health and medical history was collected, along with other demographic data, including age, sex, race, and ethnicity, as well as for weight, height, and smoking status.

Statistical analysis

Data were analyzed to describe the COVID-19 symptom profile, estimate the prevalence of individual symptoms, and identify differences between COVID-19-positive, COVID-19-negative, and untested participants. Individual variables were summarized and reported using appropriate statistical measures: mean [standard deviation (SD)] for continuous and percent (%) for categorical variables. We compared the participant characteristics at baseline and number of symptoms (0 symptoms, 1-6 symptoms, 7-9 symptoms, 10-16 symptoms) using ordered logistic regression. Nonparametric analogs were used when appropriate. Additionally, a logistic regression model was fit for each symptom to measure the association with COVID-19-positive status. Statistical significance was defined as an alpha of 0.05, with two-sided alternative hypotheses. Data analyses was conducted using Stata 16.0 (College Station, TX).

Results

As of October 31st, 2020, the CoVHORT study had enrolled a total of 1,514 participants, 245 (16.2%) of whom had lab-confirmed COVID-19. Of the remaining 1,269 participants, 339 (22.4%) participants had tested negative for COVID-19 and 930 (61.4%) had not been tested (Table 1). The participants were majority female (63.0%) and white (86.8%) and had a mean (SD) age of 47.8 (16.8) years. COVID-19-positive participants were younger (39.2 years) than COVID-19-negative participants (47.5 years), and participants who had not been tested for COVID-19 (50.1 years). COVID-19 positive participants were more likely to be Hispanic (33.5%), compared to COVID-19-negative participants (19.2%) and untested CoVHORT participants (13.8%). COVID-19-positive participants were more likely to have a body mass index (BMI) of greater than 30 kg/m^2 (34.7%) compared with COVID-19-negative participants (31.0%) and untested CoVHORT participants (23.8%). In addition, those with COVID-19 were more likely to be never smokers (93.5 vs. 86.1 and 90.3% for negative and untested participants, respectively).

Of the 245 lab-confirmed COVID-19-positive participants, the majority (85.0%) reported having experienced at least one symptom at baseline, while the remaining 38 participants (15.0%) were asymptomatic, having reported never experiencing any symptoms (Table 2). When asked to self-rate the severity of their illness on a scale of 0-10, those who reported 10-16 symptoms reported a mean (SD) severity score of 6.5 (2.0), while participants with 7-9 symptoms reported a mean

severity score of 5.5 (2.4), and participants with 1-6 symptoms reported a mean severity score of 3.3 (2.1).

Figure 1 displays the first symptom reported by COVID-19-positive study subjects who stated that they had experienced at least one symptom. The most common first symptoms were sore throat (19.0%), headache (15.5%), cough (12.6%), runny nose/cold-like symptoms (12.1%), and fatigue (12.0%). As shown in Table 3, other common symptoms that lab-confirmed COVID-19-positive participants reported at any time in their disease course included fatigue (71.8%), headache (64.5%), loss of taste or smell (57.1%), aches and pains or sore muscles (58.0%), and cough (54.3%). COVID-19-positive participants had greater odds of reporting fever, sore throat, difficulty breathing or shortness of breath, chills, diarrhea, headache, and "other symptoms" when compared to participants who tested negative for COVID-19 and participants who were never tested for COVID-19. While the magnitude of effect for these latter symptoms was smaller, all results were statistically significant. No differences between groups were observed for rash on skin, discoloration of fingers or toes, and conjunctivitis.

After adjusting for age, ethnicity, and sex, COVID-19 positive participants were more likely than negative participants to experience loss of taste and smell (OR 35.7, 95% CI 18.5-69.5), bone or nerve pain (OR 17.9, CI 6.7-47.4), vomiting (OR 10.8, CI 3.1-37.5), nausea (OR 10.5, CI 5.5-19.9), and headache (OR 8.4, CI 5.6-12.8) (Table 3). Similarly, the symptoms with the strongest association when comparing COVID-19-positive cases with the untested participants were loss of taste or smell (OR 21.1, CI 14.0-32.0), bone/nerve pain (OR 25.5, CI 12.0-54.2.0), headache (OR 11.3, CI 8.0-16.2), nausea (OR 10.9, CI 6.6-17.5), and vomiting (OR 8.7, CI 3.8-20.4).

Discussion

We determined that lab-confirmed COVID-19 cases differed in age, ethnicity, BMI and smoking status from COVID-negative participants, and untested cohort members. These same factors were associated with reported symptom severity. The most commonly reported first symptoms among COVID-19 positive participants were sore throat, followed by headache, cough, runny nose/cold-like symptoms, and fatigue. Discriminating symptoms for COVID-19-positivity included loss of taste and smell and bone or nerve pain.

Individuals identifying as Hispanic in CoVHORT constituted 33.5% of the recruited COVID-positive participants, mirroring the broader statewide case composition reported by the Arizona Department of Health Services(13). By comparison, they constituted far fewer of the lab-negative and untested groups. As discussed by Macias Gil et al.(14), the burden of COVID-19 on communities of color has been far more extreme due to extant healthcare disparities, with greater rates of hospitalizations and deaths among U.S. Hispanics as compared to whites being reported in other studies (14). Further, because publicly-available COVID-19 data by race or ethnicity may have missing values, it is critical to continue to follow up the health outcomes of this medically-vulnerable group.

Differences in disease outcomes by body size have been well-documented. In the first large study of COVID-19 patients in the United States, obesity was determined to be a major risk factor for hospitalization (3), but it remains unclear whether this finding is attributable to comorbidities that

are themselves associated with both larger body size and with severe COVID-19. In the present work, only those with a BMI of greater than 30 kg/m^2 were at increased risk for being COVID-19 positive compared to those with classified as normal weight or overweight. Disentangling the drivers of susceptibility and disease progression will require long-term follow-up in a large, diverse study population, particularly as several comorbidities, such as type 2 diabetes, are also strongly associated with larger body size. Future work from this cohort will include detailed investigations of the impact of body size on susceptibility to and recovery from COVID-19.

Another equivocal risk factor is smoking, which to date has not been clearly demonstrated to convey an increased risk for severe disease (3). In the present work, never smokers comprised 93.5% of the COVID-19 positive participants and 86.1% of the COVID-19 negative participants. This could indicate that those who do not smoke are more susceptible to infection or conversely, that those who smoke are concerned about their risk and are taking additional precautionary measures. A previous study in the United States indicated that current or former smokers were less likely to be hospitalized with COVID-19, but that former smokers were more likely to go on to develop severe disease after hospitalization, and no differences in frequency of critical illness were observed for current smokers(3). However, smoking is known to upregulate the production of the ACE2 receptor cells needed for SARS-CoV-2 to invade cells, though nicotine is known to block the ACE2 receptors(15). This paradox complicates the relationship between smoking and COVID-19. and there is significant variability in the literature. Therefore, more work is needed to assess the role of smoking in COVID-19 disease progression, and future work from CoVHORT will include a detailed analysis of different smoking modalities such as vaping or e-cigarettes, cigar, and cigarette smoking.

Several efforts have been made to identify and characterize the symptom pattern of COVID-19 to allow for more efficient and targeted screening practices, as well as to differentiate SARS-CoV-2 infection from other diseases, such as influenza (8-10, 16). However, these reports of COVID-19 symptoms have largely been confined to observational studies lacking a population-based comparison group. Because many of the symptoms reported as being associated with COVID-19 are general symptoms that could be associated with conditions such as allergies or other infectious illnesses such as influenza, there is an urgent need to evaluate the prevalence of reported symptoms of confirmed COVID-19-positive cases as compared to confirmed COVID-19-negative individuals, as well as with the prevalence of symptoms in the general population.

The results of the present study demonstrate that in southern Arizona, the most common first symptom reported by COVID-19-positive participants was sore throat, other common first symptoms of COVID-19 included headache, cough, runny nose or cold-like symptoms, and fatigue. While these are the same cluster of symptoms as reported by Larsen et al. in a large meta-analysis of more than 50,000 subjects, with data captured by the World Health Organization (WHO), the timing of appearance differed (10). Specifically, the report by Larsen concluded that the order of symptom appearance was estimated to be fever, cough, nausea, and vomiting; while in the current work, the first symptom reported by the majority of cases was sore throat, followed by headache, cough, and runny nose; only 6% of participants had fever as their first symptom. Differences in the study population, including geographic location, sex, age, timing within the

pandemic, severity of illness that prompted healthcare seeking behavior and testing, testing accessibility, and race differences across the spectrum of studies employed in the meta-analysis, may explain some of the inconsistent results for first reported symptoms.

An example of this variation in symptom reporting can be observed regarding the number of symptoms that women experienced as compared to men. Women were more likely to be classified in the category of the greatest number of symptoms than men, as were those with a BMI of greater than 30 kg/m^2, compared to those with a BMI below that threshold, although these findings were not statistically significant. A greater proportion of smokers was observed in the asymptomatic category, as compared to the any symptoms category. These findings suggest that ascertaining the type and order of COVID-19 specific symptomology may be confounded by characteristics of the participants.

With regard to overall symptom profiles, the greatest differences between laboratory-confirmed positive and negative participants were observed for loss of smell and taste and bone or nerve pain, followed by vomiting, nausea, and headache. A similar pattern was seen when comparing cases to the overall untested sample. To date, most work regarding symptoms has relied upon the frequency of symptom occurrence among cases, with little ability to ascertain the degree to which these symptoms differentiate cases from non-cases. For instance, the largest meta-analysis of COVID-19 symptomology to date included data from 24,410 cases from nine countries reported that the most common symptoms were fever (78%), cough (57%), and fatigue (31%) (8). A smaller study within the United States found that the frequency of symptoms among cases was highest for cough (84%), fever (80%), aches and pains (63%), chills (63%), and fatigue (62%)(16). In comparison, herein we found that the most common symptoms reported by cases were fatigue, headache, loss of smell or taste, cough, aches or pains, or sore muscles.

A key finding of this work is that the discrimination of COVID-19-positive symptom profiles from others requires comparison groups. General symptoms reported differ from those which may be applied to differentiate COVID-19 from other infectious diseases or conditions that are present in the underlying population. The symptoms that demonstrated the greatest difference between COVID-19-positive participants and the prevalence of symptoms among laboratory-confirmed COVID-19 negative participants or in the general population were loss of smell and taste, bone or nerve pain, headache, nausea, and fatigue.

The strengths of this study are its prospective nature, ability to capture data for laboratory-confirmed COVID-19-positive cases who have not been hospitalized, and the presence of comparison groups among both those who tested negative for COVID-19 as well as a population base drawn from throughout Arizona. These aspects allowed us to compare symptoms between cases and laboratory-confirmed uninfected individuals. However, limitations of the work must also be considered. Although we have laboratory-confirmed negative participants, we cannot know the COVID-19 status of the untested participants. It is possible that some had already been infected but were asymptomatic or exhibited few symptoms. This would likely attenuate any associations between exposure and outcomes in this study. Additionally, there may be differences in the source population for cases as compared to the laboratory-negative participants and untested participants due to the differences in recruitment strategies for these populations. For example, while postcards

were mailed to a random selection of households, it is possible Latinx participants were less likely to respond to this method than direct recruitment as cases during routine case follow-up. This could bias the association between being COVID-19-positive and Latinx away from the null. However, our race/ethnicity profile among cases is approximately similar to the overall distribution of cases throughout Arizona, suggesting a representative sample. Therefore, bias would potentially come from differential responses to other recruitment methods.

In conclusion, the findings of this analysis from the Arizona CoVHORT study show variation in several individual characteristics between COVID-19-positive participants, negative participants, and the untested population, which will be studied in future publications to assess the contributors to these observations. In addition, we found that in southern Arizona, COVID-19 positive participants most commonly reported a sore throat headache, fatigue, cough, or runny nose as the first symptom they noted. These results may aid in earlier identification of cases in the future and highlight the continued importance of addressing surveillance strategies as the pandemic continues.

Table 1. Demographic characteristics of CoVHORT participants who were laboratory-confirmed positive for COVID-19, those who were tested and were negative for COVID-19, and those without COVID-19 test results in the CoVHORT population.

		Lab-confirmed COVID-19 status	
Characteristics at study entry (baseline)	Untested participants[1] n=930	COVID-19 negative[2] n=339	COVID-19 positive[3] n=245
Age (years, mean ± sd)	50.1 (16.5)	47.5 (15.9)	39.2 (16.8)
Age (median, IQR)	51 (36,63)	48 (34,61)	38 (24, 51)
Age (range)	12-96	18-86	12-80
Sex (%)[4]			
Male	319 (34.3)	143 (42.2)	89 (36.3)
Female	605 (65.0)	193 (56.9)	155 (63.3)
Non-binary	6 (0.7)	2 (0.6)	1 (0.4)
Ethnicity (n, %)[5]			
Hispanic	128 (13.8)	65 (19.2)	82 (33.5)
Non-Hispanic	751 (80.8)	255 (75.2)	150 (61.2)
BMI (kg/m^2)			
< 18.5	8 (0.9)	6 (1.8)	6 (2.5)
18.5 – 24.9	379 (40.8)	117 (34.5)	88 (35.9)
25.0 – 29.9	306 (32.9)	105 (31.0)	64 (26.1)
30.0 – 39.9	188 (20.2)	80 (23.6)	73 (29.8)
≥ 40	33 (3.6)	25 (7.40)	12 (4.9)
Smoking status (n, %)			
Never	840 (90.3)	292 (86.1)	229 (93.5)
Occasionally	30 (3.2)	24 (7.1)	13 (5.3)
Regularly	26 (2.8)	13 (3.8)	1 (0.41)

[1]All participants in CoVHORT who do not have a laboratory-confirmed result; [2]PCR or antibody negative; [3]PCR-positive; [4]Prefer not to answer (n=1); [5]Missing data for ethnicity (n=83); BMI (n=24); smoking (n=46);

Table 2. Characteristics of COVID-19 positive study participants (n=245) by reported number of COVID-19 disease symptoms.

Characteristics at study entry (baseline)	No symptoms (n = 38)	Any symptoms (n= 207)	1-6 symptoms (n=71)	7-9 symptoms (n=71)	10-16 symptoms (n=65)	p-value[1]
Age (years, mean ± sd)	39.8 (18.18)	39.1 (16.6)	36.1 (16.7)	43.9 (17.6)	37.0 (14.2)	0.90
Days since symptoms began (mean ± sd)[2]	-	35.5 (39.4)	28.0 (33.4)	44.0 (48.2)	34.5 (33.2)	0.31
Median Days since symptoms began	-	19	13	21	19	
Sex (n, %)[3]						0.03
Female	21 (55.3)	134 (64.7)	43 (60.6)	40 (56.3)	51 (78.5)	
Male	16 (42.1)	73 (35.3)	28 (39.4)	31 (43.7)	14 (21.5)	
Ethnicity (n, %)[4]						0.51
Non-Hispanic	22 (57.9)	128 (61.8)	41 (57.8)	46 (64.8)	41 (63.1)	
Hispanic	15 (39.5)	67 (32.3)	25 (35.2)	19 (26.8)	23 (35.4)	
BMI (kg/m^2)[4]						0.08
< 18.5	2 (5.3)	4 (1.93)	1 (1.4)	-	3 (4.6)	
18.5 – 24.9	20 (52.6)	68 (32.9)	30 (42.3)	18 (25.4)	20 (30.8)	
25.0 – 29.9	8 (21.1)	56 (27.1)	19 (26.8)	19 (26.8)	18 (27.7)	
30.0 – ≥40	8 (21.1)	77 (37.2)	21 (29.6)	34 (47.9)	22 (33.9)	
Smoking Status[4]						0.70
Never	37 (97.4)	192 (92.8)	64 (90.1)	66 (93.0)	62 (95.4)	
Occasionally or Regularly	1 (2.6)	13 (6.3)	6 (8.5)	5 (7.0)	2 (31)	

[1]P-values calculated using ordered logistic regression. [2]Number of days between start of symptoms and survey completion, missing data for number of days between start of symptoms and survey completion (n=4); [3]Non-binary (n=1). [4]Missing data for ethnicity (n=13); BMI (n=2); smoking status (n=2).

Table 3. Symptom characteristics of CoVHORT participants by case status, adjusted by age, sex, and ethnicity.

Reported symptoms at study entry (baseline)	COVID-19 positive[1] n=245	Untested participants[2] n=930	COVID-19 negative[3] n=339	Positive vs Untested	Positive vs Negative
	n (%)	n (%)	n (%)	OR (95% C I)	OR (95% CI)
Fatigue	176 (71.8)	170 (18.3)	87 (25.7)	11.4 (8.0, 16.3)	8.3 (5.5, 12.5)
Headache	158 (64.5)	120 (12.9)	60 (17.7)	11.3 (8.0, 16.2)	8.4 (5.6, 12.8)
Aches and pains or sore muscles	142 (58.0)	130 (14.0)	66 (19.5)	8.5 (6.0, 11.9)	6.4 (4.3, 9.7)
Loss of smell/taste	140 (57.1)	48 (5.2)	12 (3.5)	21.1 (14.0, 32.0)	35.7 (18.4, 69.5)
Cough	133 (54.3)	156 (16.8)	75 (22.1)	5.5 (3.9, 7.6)	4.0 (2.7, 5.9)
Fever	122 (49.8)	120 (12.9)	61 (18.0)	6.7 (4.8, 9.5)	5.2 (3.5, 7.9)
Runny nose/cold-like symptoms	115 (46.9)	105 (11.3)	43 (12.7)	6.4 (4.5, 9.1)	6.3 (4.0, 9.8)
Chills	104 (42.5)	89 (9.6)	39 (11.5)	6.9 (4.7, 9.9)	6.8 (4.3, 10.7)
Sore throat	101 (41.2)	123 (13.2)	61 (18.0)	3.9 (2.7, 5.5)	2.9 (2.0, 4.4)
Difficulty breathing or shortness of breath	92 (37.6)	94 (10.1)	48 (14.2)	5.4 (3.7, 7.8)	4.2 (2.7, 6.6)
Diarrhea	81 (33.0)	50 (5.4)	19 (5.6)	8.4 (5.5, 12.9)	9.0 (5.1, 16.1)
Nausea	79 (32.2)	34 (3.7)	14 (4.1)	10.9 (6.6, 17.5)	10.5 (5.5, 19.9)
Chest pain or pressure	74 (30.2)	61 (6.6)	41 (12.1)	6.4 (4.2, 9.7)	3.4 (2.1, 5.3)
Bone pain/nerve pain	47 (19.2)	10 (1.1)	5 (1.5)	25.5 (12.0, 54.2)	17.9 (6.7, 47.4)
Vomiting	24 (9.8)	9 (1.0)	3 (0.9)	8.7 (3.8, 20.4)	10.8 (3.1, 37.5)
Other	17 (6.9)	9 (0.97)	3 (0.9)	8.4 (3.3, 21.4)	8.1 (2.2, 29.2)
Rash on skin	15 (6.1)	8 (0.9)	4 (1.2)	10.9 (4.1, 28.7)	7.9 (0.86, 73.4)
Discoloration of fingers/toes	5 (2.0)	3 (0.3)	1 (0.3)	7.1 (1.6, 31.9)	7.0 (0.82, 60.7)
Loss of speech or movement	2 (0.8)	1 (0.1)	-	7.0 (0.5, 91.8)	-
Conjunctivitis	2 (0.8)	7 (0.8)	2 (0.6)	1.21 (0.22, 6.6)	1.2 (0.16, 8.6)

[1]PCR-positive cases; [2]participants in CoVHORT who do not have a laboratory-confirmed result; [3]PCR or antibody negative.

Figure 1. First symptom reported by participants who are laboratory-confirmed positive COVID-19 cases.

Ethics Statement: This study involving human participants was reviewed and approved by the Institutional Review Board of the University of Arizona Human Subjects Protection Program (#2003521636A00). Written informed consent to participate in this study was provided by the participants or the participants' legal guardian/next of kin.

Funding Statement: This work was supported by the BIO5 Institute at The University of Arizona. Grant Number: N/A.

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Data Sharing: No additional data available.

Competing Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential competing interest.

Contributorship statement: KP-B, LF, MJ, MB, ZC, YK, KE, EJ, PG-F, SC, LK and FC-M conceptualized the study and developed the initial study protocol. SK, KH, CC, EA, JH, EJ, CK and KP-B participated in the design of the protocol, and the drafting and reviewing of the manuscript. All authors critically reviewed the draft of the manuscript and approved the final version. All authors meet the criteria for authorship as developed by the International Committee for Medical Journal Editors.

References

1. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. New England Journal of Medicine. 2020;382(8):727-33.
2. Guan WJ, Ni ZY, Hu Y, Liang WH, Ou CQ, He JX, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. N Engl J Med. 2020;382(18):1708-20.
3. Petrilli CM, Jones SA, Yang J, Rajagopalan H, O'Donnell L, Chernyak Y, et al. Factors associated with hospital admission and critical illness among 5279 people with coronavirus disease 2019 in New York City: prospective cohort study. BMJ. 2020;369:m1966.
4. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet. 2020;395(10229):1054-62.
5. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. Lancet. 2020;395(10223):507-13.
6. Cai H. Sex difference and smoking predisposition in patients with COVID-19. Lancet Respir Med. 2020;8(4):e20.
7. Hygiene NDoHaM. Age-adjusted rates of lab confirmed COVID-19 non-hospitalized cases, estimated non-fatal hospitalized cases, and persons known to have died per 100,000 by race/ethnicity group. Albany, NY2020.
8. Grant MC, Geoghegan L, Arbyn M, Mohammed Z, McGuinness L, Clarke EL, et al. The prevalence of symptoms in 24,410 adults infected by the novel coronavirus (SARS-CoV-2; COVID-19): A systematic review and meta-analysis of 148 studies from 9 countries. PloS one. 2020;15(6):e0234765.
9. Lechien JR, Chiesa-Estomba CM, Place S, Van Laethem Y, Cabaraux P, Mat Q, et al. Clinical and epidemiological characteristics of 1420 European patients with mild-to-moderate coronavirus disease 2019. J Intern Med. 2020;288(3):335-44.
10. Larsen JR, Martin MR, Martin JD, Kuhn P, Hicks JB. Modeling the Onset of Symptoms of COVID-19. Front Public Health. 2020;8:473.
11. Catalfamo CJ, Heslin KM, Shilen A, Khan SM, Hunsaker JR, Austhof E, et al. Design of the Arizona CoVHORT: A Population-Based COVID-19 Cohort. Frontiers in Public Health. 2021;9(21).
12. Dillman D, Smyth J, Christian L. Internet, Mail, and Mixed-Mode Surveys: The Tailored Design Method. Hoboken, NJ: Wiley & Sons; 2009.
13. Services ADoH. COVID-19 Data Dashboard Phoenix, Arizona: Arizona Department of Health Services; 2021 [Available from: https://www.azdhs.gov/preparedness/epidemiology-disease-control/infectious-disease-epidemiology/covid-19/dashboards/index.php.
14. Macias Gil R, Marcelin JR, Zuniga-Blanco B, Marquez C, Mathew T, Piggott DA. COVID-19 Pandemic: Disparate Health Impact on the Hispanic/Latinx Population in the United States. J Infect Dis. 2020;222(10):1592-5.
15. Smith JC, Sausville EL, Girish V, Yuan ML, Vasudevan A, John KM, et al. Cigarette Smoke Exposure and Inflammatory Signaling Increase the Expression of the SARS-CoV-2 Receptor ACE2 in the Respiratory Tract. Dev Cell. 2020;53(5):514-29 e3.

16. Burke RM, Killerby ME, Newton S, Ashworth CE, Berns AL, Brennan S, et al. Symptom Profiles of a Convenience Sample of Patients with COVID-19 - United States, January-April 2020. MMWR Morb Mortal Wkly Rep. 2020;69(28):904-8.

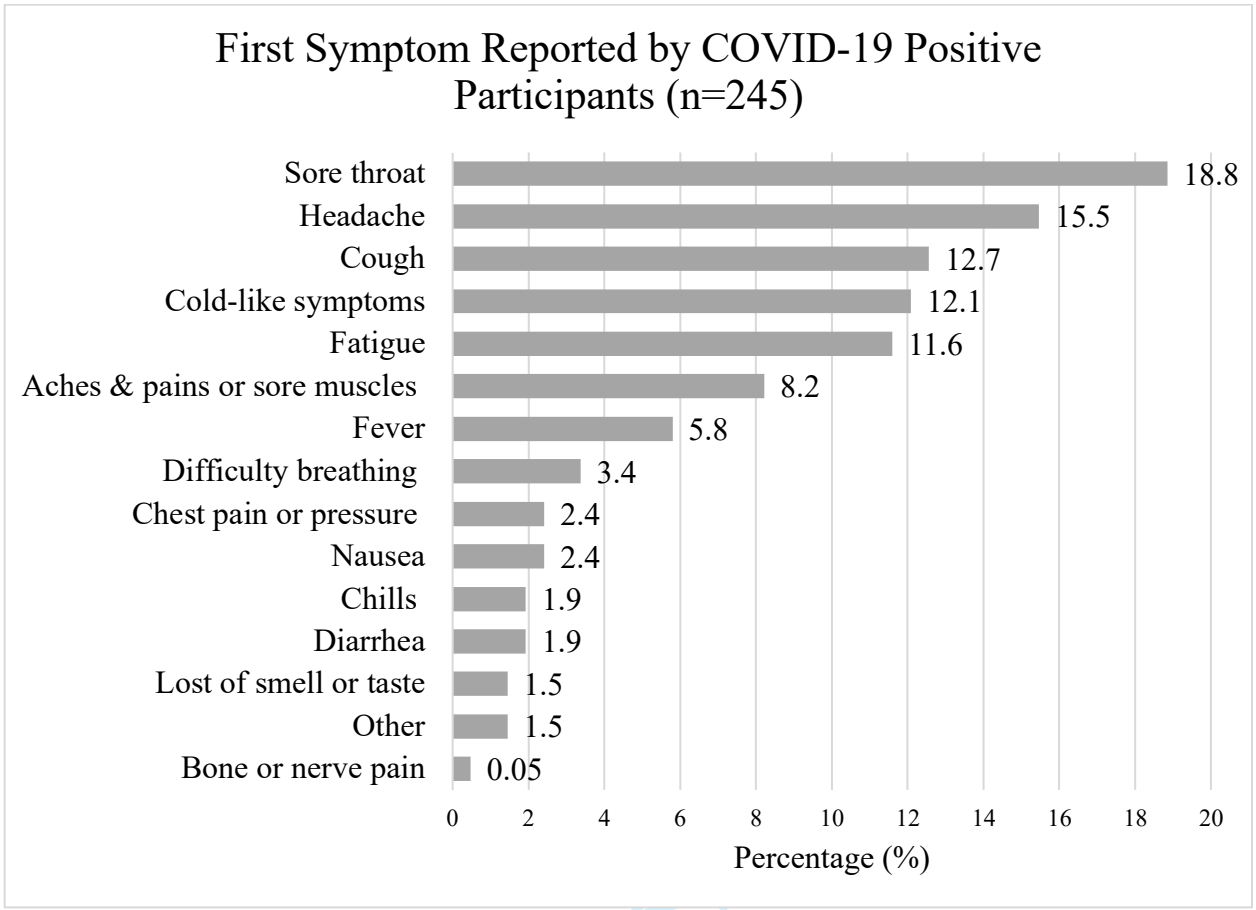

Figure 1. First symptom reported by participants who are laboratory-confirmed positive COVID-19 cases.

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract (*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	1,3
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4,5
Objectives	3	State specific objectives, including any prespecified hypotheses	4,5
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up (*b*) For matched studies, give matching criteria and number of exposed and unexposed	5
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	5
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	5
Bias	9	Describe any efforts to address potential sources of bias	5,9
Study size	10	Explain how the study size was arrived at	5
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	5
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding (*b*) Describe any methods used to examine subgroups and interactions (*c*) Explain how missing data were addressed (*d*) If applicable, explain how loss to follow-up was addressed (*e*) Describe any sensitivity analyses	6
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed (b) Give reasons for non-participation at each stage (c) Consider use of a flow diagram	6
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders (b) Indicate number of participants with missing data for each variable of interest (c) Summarise follow-up time (eg, average and total amount)	6
Outcome data	15*	Report numbers of outcome events or summary measures over time	6.7

Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included (*b*) Report category boundaries when continuous variables were categorized (*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	6,7
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	6,7
Discussion			
Key results	18	Summarise key results with reference to study objectives	6,7
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	9
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	7-10
Generalisability	21	Discuss the generalisability (external validity) of the study results	7-10
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	10

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.

BMJ Open

Elucidating symptoms of COVID-19 illness in the Arizona CoVHORT, a longitudinal cohort study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-053403.R1
Article Type:	Original research
Date Submitted by the Author:	09-Nov-2021
Complete List of Authors:	Khan, Sana; The University of Arizona, Department of Epidemiology and Biostatistics Farland, Leslie V.; The University of Arizona, Department of Epidemiology and Biostatistics; The University of Arizona Cancer Center Catalfamo, Collin; The University of Arizona, Department of Epidemiology and Biostatistics Austhof, Erika; The University of Arizona, Department of Epidemiology and Biostatistics Bell, Melanie L.; The University of Arizona, Department of Epidemiology and Biostatistics Chen, Zhao; The University of Arizona, Department of Epidemiology and Biostatistics Cordova-Marks, Felina; The University of Arizona Cancer Center; The University of Arizona, Department of Health Promotion Sciences Ernst, Kacey; The University of Arizona, Department of Epidemiology and Biostatistics, College of Public Health Garcia-Filion, Pamela; The University of Arizona, Department of Biomedical Informatics Heslin, Kelly M.; The University of Arizona, Department of Epidemiology and Biostatistics Hoskinson, Joshua; The University of Arizona Cancer Center Jehn, Megan; Arizona State University Joseph, Emily C.S.; The University of Arizona Cancer Center Kelley, Connor; The University of Arizona, Department of Epidemiology and Biostatistics Klimentidis, Yann; The University of Arizona, Department of Epidemiology and Biostatistics Russo Carroll, Stephanie; The University of Arizona, Department of Community, Environment, and Policy; The University of Arizona, Native Nations Institute at the Udall Center for Studies in Public Policy Kohler, Lindsay; The University of Arizona, Department of Epidemiology and Biostatistics; The University of Arizona, Department of Health Promotion Sciences Pogreba-Brown, Kristen; The University of Arizona, Department of Epidemiology and Biostatistics Jacobs, Elizabeth; The University of Arizona, Department of Epidemiology and Biostatistics; The University of Arizona Cancer Center
Primary Subject Heading:	Infectious diseases

Secondary Subject Heading:	Epidemiology
Keywords:	COVID-19, INFECTIOUS DISEASES, EPIDEMIOLOGY

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Elucidating symptoms of COVID-19 illness in the Arizona CoVHORT, a longitudinal cohort study

Sana M. Khan[1]

Leslie V. Farland[1,2]

Collin J. Catalfamo[1]

Erika Austhof[1]

Melanie L. Bell[1]

Zhao Chen[1]

Felina Cordova-Marks[2,3]

Kacey C. Ernst[1]

Pamela Garcia-Filion[4]

Kelly M. Heslin[1]

Joshua Hoskinson[2]

Megan L Jehn[5]

Emily C.S. Joseph[2]

Connor P. Kelley[1]

Yann C. Klimentidis[1]

Stephanie Russo Carroll[6,7]

Lindsay N. Kohler[1,3]

Kristen Pogreba-Brown[1]

Elizabeth T. Jacobs[1,2]

[1]Department of Epidemiology and Biostatistics, Mel and Enid Zuckerman College of Public Health, The University of Arizona, Tucson, AZ, United States; [2]University of Arizona Cancer Center, Tucson, AZ, United States, [3]Department of Health Promotion Sciences, Mel and Enid Zuckerman College of Public Health, The University of Arizona, Tucson, AZ, United States; [4]Department of Biomedical Informatics, College of Medicine–Phoenix, The University of Arizona, Phoenix, AZ, United States, [5]School of Human Evolution and Social Change, Arizona State University, Tempe, AZ, United States; [6]Department of Community, Environment and Policy, Mel and Enid Zuckerman College of Public Health, The University of Arizona, Tucson, AZ, United States; [7] Native Nations Institute at the Udall Center for Studies in Public Policy, University of Arizona, Tucson, Arizona.

Corresponding Author

Sana M. Khan, MPH
Doctoral Student, Epidemiology
University of Arizona
Mel and Enid Zuckerman College of Public Health
sanakhan@email.arizona.edu

Keywords: SARS-CoV-2, COVID-19, symptoms, prospective cohort

Abstract

Objective: To elucidate the symptoms of laboratory-confirmed COVID-19 cases as compared to laboratory-confirmed negative individuals and to the untested general population among all participants who reported symptoms within a large, prospective cohort study. **Setting and Design:** This work was conducted within the framework of The Arizona CoVHORT, a longitudinal prospective cohort study conducted among Arizona residents. **Participants:** Eligible participants were any individual living in Arizona and were recruited from across Arizona via COVID-19 case investigations, participation in testing studies, and a postcard mailing effort. **Primary and Secondary Outcome Measures:** The primary outcome measure was a comparison of the type and frequency of symptoms between COVID-19 positive cases, tested but negative individuals, and the general untested population who reported experiencing symptoms consistent with COVID-19. **Results:** Of the 1,335 laboratory-confirmed COVID-19 cases, 180 (13.5%) reported having no symptoms. Of those that did report symptoms, the most commonly reported were fatigue (82.2%), headache (74.6%), aches, pains, or sore muscles (66.3%), loss of taste or smell (62.8,) and cough (61.9%). In adjusted logistic regression models, COVID-19 positive participants were more likely than negative participants to experience loss of taste and smell (OR 12.1; 95% CI 9.6-15.2); bone or nerve pain (OR 3.0; 95% CI 2.2 - 4.1), headache (OR: 2.6; 95% CI 2.2-3.2), nausea (OR: 2.4; 95% CI 1.9-3.1), or diarrhea (OR: 2.1; 95% CI 1.7-2.6). Fatigue (82.9) and headache (74.9) had the highest sensitivities among symptoms, while loss of taste or smell (87.2) and bone or nerve pain (92.9) had the high specificities among significant symptoms associated with COVID-19. **Conclusion:** When comparing confirmed COVID-19 cases with either confirmed negative or untested participants, the pattern of symptoms that discriminates SARS-CoV-2 infection from those arising from other potential circulating pathogens may differ from general reports of symptoms among cases alone.

Strengths and limitations of this study:

- The findings of this study will aid in the identification of symptoms that differentiate COVID-19 from other circulating infections or conditions, which is useful for resource-limited settings where diagnostic testing is limited.
- To our knowledge, no prior research has compared the prevalence of non-specific symptoms such as headache, fever, and runny nose between confirmed COVID-19-positive cases, confirmed COVID-19-negative cases, and a general, untested comparison groups.
- We cannot know the COVID-19 status of the untested participants; it is possible that some had already been infected but were asymptomatic or exhibited few symptoms.

Introduction

In late 2019, the novel coronavirus SARS-CoV-2 was first recognized in China among patients who presented with pneumonia and the first scientific report appeared shortly thereafter [1]. On March 11th, 2020, the World Health Organization declared COVID-19 a pandemic. The pathogen has had multiple impacts on individual and societal wellbeing arising from both biological effects of the virus and policy-based mitigation. The majority of those infected with acute COVID-19 will go on to recover, though approximately 10-20% of COVID-19 patients overall will develop a severe case of disease, and may suffer from stroke, pneumonia, or acute respiratory distress syndrome (ARDS) and require intensive care and ventilation [2, 3].

Individuals are likely be most infectious during the early phases of the disease, when symptoms may be comparatively mild; therefore, it is important to elucidate the reported symptom patterns of COVID-19 patients compared to both laboratory-confirmed negative individuals and population-based controls. Several risk factors have been associated with disease susceptibility and severity including increasing age [4], male sex [2, 5, 6], and current or former smoking [3], which may also affect symptomology. Further, important differences in disease incidence and severity by race and ethnicity have emerged, with Native Americans, African Americans, and Latinos having higher COVID-19 prevalence, hospitalization, and mortality rates compared to non-Hispanic whites [7]. It is presently not well known if reports of symptoms or symptom patterns vary by these factors as well.

A recent meta-analysis of over 24,000 patients across nine countries reported on COVID-19 symptom presentation. In this work, the most commonly reported symptoms among people with COVID-19 were fever (78% of COVID-19 patients reporting), cough (57%), and fatigue (31%) [8]. Additionally, a systematic review published in February 2021 aimed to assess the diagnostic accuracy of symptoms associated with COVID-19; this review identified 44 studies which in total included over 26,000 participants. The review found that among 84 symptoms, cough and fever had high sensitivities and could be used as a prompt for further COVID-19 testing. However, a limitation of the review article is potential selection bias due to their sample being selected from primarily clinical settings [9]. Additional work examining symptoms in an unselected population is necessary to determine the syndromic presentation of COVID-19 in the general population. Another study conducted among European patients (n=1420) with mild or moderate COVID-19 found that the most frequently reported symptoms were headache (70%), loss of smell (70%), and obstruction of the nasal passages (68%) [10]. The authors of a separate study, the objective of which was to develop a better symptom modeling algorithm to aid targeted testing, concluded that fever and cough should be used as the key symptoms for rapid COVID-19 screening given their high sensitivity [11]. However, a major limitation of studies conducted to date is the lack of comparison of patient-reported symptoms to those of uninfected individuals. To our knowledge, no prior research has compared the prevalence of non-specific symptoms such as headache, fever, and runny nose between confirmed COVID-19-positive cases, confirmed COVID-19-negative cases, and population-based comparison groups.

Since COVID-19 community transmission began, Arizona has experienced multiple, severe, COVID-19 surges, with more than 1.1 million infections and 21,000 COVID-19-related deaths as

of October 2021. To address this epidemiological challenge, in May 2020, we initiated a large, prospective cohort in Arizona of racially- and ethnically- diverse residents in order to rigorously investigate factors contributing to variability in natural COVID-19 disease history including incidence, progression, resolution, and chronic outcomes of infection [12]. This COVID-19 cohort, dubbed The Arizona CoVHORT, provides a rich data source for multiple areas of inquiry related to the pandemic. The objective of the present work was to determine which symptoms were reported with the greatest frequency among participants who tested positive for COVID-19 as compared to participants who tested negative for COVID-19 and untested participants, while controlling for potential confounders such as age, ethnicity, sex, BMI, and smoking status. The findings of this paper will aid in the identification of symptoms that differentiate COVID-19 from other circulating infections or conditions, such as allergies.

Materials and Methods

Study Participants

The overall goal of the CoVHORT is to continuously enroll Arizonans into a cohort study to track both the acute and long-term phases of infection with SARS-CoV-2. The present analysis includes data through October 1st, 2021. Several recruitment methods were employed, which have been described in detail previously [12]. Briefly, the primary sources of recruitment have been through case investigations in a partnership with the Arizona Department of Health Services and other research studies and testing sites at the University of Arizona and Arizona State University, both of which have allowed for inclusion of laboratory-confirmed COVID-19 positive and negative participants. By October 1st, 2021, a total of 493 COVID-19-positive participants had been recruited through health department case investigations and 901 through our partnerships with studies and testing sites in Arizona.

A comprehensive mailing list was purchased that provides information on 2.2 million residents in Arizona. To recruit the population-based comparison group, a total of 17,500 postcards were mailed to a simple random sample of Pima County, Arizona residents in July 2020. Consistent with the Dillman method to maximize participation and minimize bias [13], three phased mailings of recruitment postcards occurred every two weeks. Participant-provided information from returned surveys was used to exclude those who had already enrolled from subsequent phases of the mailing campaign. Each list was screened prior to each mailing to reduce the number of undeliverable postcards. We have completed all three phases of the mailing campaign in Pima County, with 17,294 postcards delivered in the first phase, 17,147 in the second phase, and another 17,081 in the third phase. Method of recruitment is recorded for all participants allowing sensitivity analyses to be conducted within subgroups.

Patient and Public Involvement

We encourage active participation from members of The Arizona CoVHORT. The public and members of the cohort are invited to webinars where they are able to provide input, ask questions, and speak with the projects' principal investigators. We regularly revisit our survey instruments to ensure they are reflecting feedback from participants and are centering their experiences and

priorities. Study findings are disseminated at our study website (covhort.arizona.edu), along with a regularly updated participant dashboard containing descriptive data of the cohort population.

Survey Instruments

All participants included in the CoVHORT were sent identical structured electronic questionnaires at upon study entry, regardless of COVID-19 status. All participants were first asked, "Since January 1, have you experienced a sudden illness that led you to believe you had COVID19?" If they answered "yes", all participants, were asked to indicate which symptoms they had experienced since January 2020 from a list based upon prior reports in the literature, as well as through an open-text field. Participants who respond "no" are not asked about symptomology and were not included in this analysis. Regardless of symptom status, all participants were then asked if they had tested for the virus that causes COVID-19 with a nasal swab, throat swab, or saliva test since January 2020. Participants were classified as untested, positive, or negative based on their results (Table 1). Information regarding health and medical history was collected, along with other demographic data, including age, sex, race, and ethnicity, as well as for weight, height, and smoking status. From these data, we calculated body mass index as (kg/m^2), and categorized participants as having a BMI of <25, ≥ 25-29.9, and ≥ 30, to aid in clinical interpretation, as well as reported BMI as a continuous variable (Table 2).

Statistical analysis

Data were analyzed to describe the COVID-19 symptoms, estimate the prevalence of individual symptoms, and identify differences among COVID-19-positive cases compared to COVID-19-negative individuals and untested participants. Individual variables were summarized and reported using appropriate statistical measures: mean [standard deviation (SD)] for continuous and percent (%) for categorical variables. Among those who tested positive for COVID-19, we compared the participant characteristics upon study entry and number of symptoms (0 symptoms, 1-6 symptoms, 7-9 symptoms, 10-16 symptoms) using ordered logistic regression and report p-values to explore factors associated with increasing severity. A logistic regression model was fit for each symptom to measure the association, as measured by odds ratios and 95% confidence intervals, with COVID-19-positive status after adjusting for age, sex, ethnicity, BMI, and smoking status. Additionally, we included sensitivity and specificity analysis for each individual symptom (Supplemental Table 1). Statistical significance was defined as $p < 0.05$, with two-sided tests. Data analyses was conducted using Stata 16.0 (College Station, TX).

Results

As of October 1st, 2021, the Arizona CoVHORT study had enrolled a total of 7,012 participants, 2,373 (33.8%) of whom reported symptoms associated with COVID-19 since January 2021. Of these participants, 1,335 (56.3%) had lab-confirmed positive COVID-19 result, 930 (39.2%) had a lab-confirmed COVID-19 negative result, and 288 (12.1%) were untested (Table 1). The participants were majority female (70.4%) and white (89.4%) and had a mean (SD) age of 44.5 (15.3) years. COVID-19-positive participants were younger (43.9 years) than COVID-19-negative participants (44.8 years), and participants who had not been tested for COVID-19 (46.8 years). COVID-19 positive participants were more likely to be Hispanic (22.2%), compared to COVID-

19-negative participants (14.0%) and untested CoVHORT participants (16.0%). COVID-19-positive participants were more likely to have a body mass index (BMI) of greater than 30 kg/m^2 (34.3%) compared with COVID-19-negative participants (29.5%) and untested CoVHORT participants (27.2%). Of the 1,335 lab-confirmed COVID-19-positive participants, the majority (86.5%) reported having experienced at least one symptom after diagnosis, while the remaining 180 participants (13.5%) were asymptomatic, having reported never experiencing any symptoms (Table 2). When asked to self-rate the severity of their illness on a scale of 0-10, those who reported 10-18 symptoms reported a mean (SD) severity score of 6.8 (1.9), while participants with 7-9 symptoms reported a mean severity score of 5.6 (2.1), and participants with 1-6 symptoms reported a mean severity score of 3.6 (2.3) (Table 2). We assessed days since symptom onset and days since test date with the survey completion date and found no significant difference between symptom groups (Table 2).

As shown in Table 3, other common symptoms that lab-confirmed COVID-19-positive participants reported at any time in their disease course included fatigue (82.9%), headache (74.6%), loss of taste or smell (62.8%), aches and pains or sore muscles (66.3%), and cough (61.9%). COVID-19-positive participants had greater odds of reporting loss of taste or smell, bone or nerve pain, headache, nausea, and cold-like symptoms when compared to participants who tested negative for COVID-19 and participants who were never tested for COVID-19. While the magnitude of effect for these latter symptoms was smaller, all results were statistically significant. No differences between groups were observed for cough, fever, sore throat, loss of speech or movement, discoloration of fingers or toes, and conjunctivitis. After adjusting for age, ethnicity, sex, BMI, and smoking status, COVID-19 positive participants were more likely than negative participants to experience loss of taste and smell (OR 12.1, 95% CI 9.6-15.2), bone or nerve pain (OR 3.0, CI 2.2-4.1), headache (OR 2.6, CI 2.2-3.2), nausea (OR 2.4, CI 1.9-3.1), and diarrhea (OR 2.1, CI 1.7-2.6) (Table 3). Similarly, the symptoms with the strongest association when comparing COVID-19-positive cases with the untested participants were loss of taste or smell (OR 5.8, CI 4.2-7.9), bone/nerve pain (OR 2.9, CI 1.8-4.6), headache (OR 2.1, CI 1.6-2.7), nausea (OR 1.7, CI 1.2-2.5), and cold-like symptoms (OR 1.5, CI 1.1-2.0). Fatigue (82.9), headache (74.6), and aches and pains or sore muscles (66.3) were shown to have the highest sensitivities among symptoms, while loss of taste or smell (87.2) and bone or nerve pain (92.9) had high specificity among the significant symptoms (Supplemental Table 1).

Discussion

We assessed the type and frequency of symptoms between COVID-19 positive cases, tested but negative individuals, and the general untested population who reported experiencing symptoms consistent with COVID-19. We determined that lab-confirmed COVID-19 cases differed in age, ethnicity, BMI from COVID-negative participants, and untested cohort members. These same factors were associated with reported symptom severity. The most commonly reported first symptoms among COVID-19 positive participants were sore throat, followed by headache, cough, runny nose/cold-like symptoms, and fatigue. Discriminating symptoms for COVID-19-positivity included loss of taste and smell and bone or nerve pain as demonstrated by specificity analyses;

while fatigue, headache, and aches and pains or sore muscles were shown to have the highest sensitivities among symptoms.

Individuals identifying as Hispanic in CoVHORT constituted 33.5% of the recruited COVID-positive participants, mirroring the broader statewide case composition reported by the Arizona Department of Health Services [14]. By comparison, they constituted far fewer of the lab-negative and untested groups. As discussed by Macias Gil et al. [15], the burden of COVID-19 on communities of color has been far more extreme due to extant healthcare disparities, with greater rates of hospitalizations and deaths among U.S. Hispanics as compared to whites being reported in other studies [15]. Further, because publicly-available COVID-19 data by race or ethnicity may have missing values, it is critical to continue to follow up the health outcomes of this medically-vulnerable group.

Differences in disease outcomes by body size have been well-documented. In the first large study of COVID-19 patients in the United States, obesity was determined to be a major risk factor for hospitalization [3], but it remains unclear whether this finding is attributable to comorbidities that are themselves associated with both larger body size and with severe COVID-19. In the present work, only those with a BMI of greater than 30 kg/m^2 were at increased risk for being COVID-19 positive compared to those with classified as normal weight or overweight. Disentangling the drivers of susceptibility and disease progression will require long-term follow-up in a large, diverse study population, particularly as several comorbidities, such as type 2 diabetes, are also strongly associated with larger body size. Future work from this cohort will include detailed investigations of the impact of body size on susceptibility to and recovery from COVID-19.

Another equivocal risk factor is smoking, which to date has not been clearly demonstrated to convey an increased risk for severe disease [3]. In the present work, there was no difference in COVID-19 test status by smoking status. A previous study in the United States indicated that current or former smokers were less likely to be hospitalized with COVID-19, but that former smokers were more likely to go on to develop severe disease after hospitalization, and no differences in frequency of critical illness were observed for current smokers [3]. However, smoking is known to upregulate the production of the ACE2 receptor cells needed for SARS-CoV-2 to invade cells, though nicotine is known to block the ACE2 receptors [16]. This paradox complicates the relationship between smoking and COVID-19. and there is significant variability in the literature. Therefore, more work is needed to assess the role of smoking in COVID-19 disease progression, and future work from CoVHORT will include a detailed analysis of different smoking modalities such as vaping or e-cigarettes, cigar, and cigarette smoking.

Several efforts have been made to identify and characterize the symptoms associated with COVID-19 to allow for more efficient and targeted screening practices, as well as to differentiate SARS-CoV-2 infection from other diseases, such as influenza [8-10, 17]. However, these reports of COVID-19 symptoms have largely been confined to hospitalized or outpatient patient population and are lacking a symptomatic COVID-19 negative comparison group. Because many of the symptoms reported as being associated with COVID-19 are general symptoms that could be associated with conditions such as allergies or other infectious illnesses such as influenza, there is an urgent need to evaluate the prevalence of reported symptoms of confirmed COVID-19-positive

cases as compared to confirmed COVID-19-negative individuals, as well as with the prevalence of symptoms in the general population.

The results of the present study demonstrate that in southern Arizona, the most common first symptom reported by COVID-19-positive participants was sore throat, other common first symptoms of COVID-19 included headache, cough, runny nose or cold-like symptoms, and fatigue. While these are the same cluster of symptoms as reported by Larsen et al. in a large meta-analysis of more than 50,000 subjects, with data captured by the World Health Organization (WHO), the timing of appearance differed [11]. Specifically, the report by Larsen concluded that the order of symptom appearance was estimated to be fever, cough, nausea, and vomiting; while in the current work, the first symptom reported by the majority of cases was sore throat, followed by headache, cough, and runny nose; only 6% of participants had fever as their first symptom. Differences in the study population, including geographic location, sex, age, timing within the pandemic, severity of illness that prompted healthcare seeking behavior and testing, testing accessibility, and race differences across the spectrum of studies employed in the meta-analysis, may explain some of the inconsistent results for first reported symptoms.

An example of this variation in symptom reporting can be observed regarding the number of symptoms that women experienced as compared to men. Women were more likely to be classified in the category of the greatest number of symptoms than men, as were those with a BMI of greater than 30 kg/m^2, compared to those with a BMI below that threshold, although these findings were not statistically significant. A greater proportion of smokers was observed in the asymptomatic category, as compared to the any symptoms category. These findings suggest that ascertaining the type and order of COVID-19 specific symptomology may be confounded by characteristics of the participants.

With regard to overall COVID-19 symptoms, the greatest differences between laboratory-confirmed positive and negative participants were observed for loss of smell and taste and bone or nerve pain, followed by vomiting, nausea, and headache. A similar pattern was seen when comparing cases to the overall untested sample. To date, most work regarding symptoms has relied upon the frequency of symptom occurrence among cases, with little ability to ascertain the degree to which these symptoms differentiate cases from non-cases. For instance, the largest meta-analysis of COVID-19 symptomology to date included data from 24,410 cases from nine countries reported that the most common symptoms were fever (78%), cough (57%), and fatigue (31%) [8]. A smaller study within the United States found that the frequency of symptoms among cases was highest for cough (84%), fever (80%), aches and pains (63%), chills (63%), and fatigue (62%) [17]. In comparison, herein we found that the most common symptoms reported by cases were fatigue, headache, loss of smell or taste, cough, aches or pains, or sore muscles.

A key finding of this work is that the discrimination of COVID-19-positive symptoms from others requires comparison groups. General symptoms reported differ from those which may be applied to differentiate COVID-19 from other infectious diseases or conditions that are present in the underlying population. The symptoms that demonstrated the greatest difference between COVID-19-positive participants and the prevalence of symptoms among laboratory-confirmed COVID-19

negative participants or in the general population were loss of smell and taste, bone or nerve pain, headache, nausea, and fatigue.

The strengths of this study are its prospective nature, ability to capture data for laboratory-confirmed COVID-19-positive cases who have not been hospitalized, and the presence of comparison groups among both those who tested negative for COVID-19 as well as a population base drawn from throughout Arizona. These aspects allowed us to compare symptoms between cases and laboratory-confirmed uninfected individuals. However, limitations of the work must also be considered. First and foremost, while we are able to recruit participants via follow up to COVID-19 testing, participants' test results and symptoms are self-reported. Furthermore, although we have self-reported, laboratory-confirmed negative participants, we cannot know the COVID-19 status of the untested participants. It is possible that some had already been infected but were asymptomatic or exhibited few symptoms. This would likely attenuate any associations between exposure and outcomes in this study. It is also important to acknowledge that participants who indicated not experiencing symptoms that led them to believe they had COVID-19, regardless of testing status, were not asked to indicate which symptoms they had experienced. These participants were not able to directly indicate that they experienced no symptoms from the provided list; however, because the majority of CoVHORT participants who undergo testing for COVID-19 enroll at a time point after receiving their test results, we believe that the likelihood that participants who indicate not experiencing symptoms actually experienced symptoms from our list offered to participants who indicate experiencing symptoms is low. Additionally, there may be differences in the source population for cases as compared to the laboratory-negative participants and untested participants due to the differences in recruitment strategies for these populations. For example, while postcards were mailed to a random selection of households, it is possible Latinx participants were less likely to respond to this method than direct recruitment as cases during routine case follow-up. This could bias the association between being COVID-19-positive and Latinx away from the null. However, our race/ethnicity profile among cases is approximately similar to the overall distribution of cases throughout Arizona, suggesting a representative sample. Therefore, bias would potentially come from differential responses to other recruitment methods. This was an exploratory study, with a large number of statistical tests, and therefore care should be taken when considering p-values.

In conclusion, the findings of this analysis from the Arizona CoVHORT study show variation in several individual characteristics between COVID-19-positive participants, negative participants, and the untested population, which will be studied in future publications to assess the contributors to these observations. In addition, we found that in southern Arizona, COVID-19 positive participants most commonly reported a sore throat headache, fatigue, cough, or runny nose as the first symptom they noted. These results may aid in earlier identification of cases in the future and highlight the continued importance of addressing surveillance strategies as the pandemic continues.

Table 1. Demographic characteristics of CoVHORT participants who reported symptoms and were laboratory-confirmed positive for COVID-19, those who were tested and were negative for COVID-19, and those without COVID-19 test results in the CoVHORT population.

		Lab-confirmed COVID-19 status	
Characteristics at study entry	Untested participants[1,4] n= 288	COVID-19 negative[2,5] n= 930	COVID-19 positive[3,6] n= 1,335
Age [years, mean (SD)]	46.8 (14.2)	44.8 (14.0)	43.9 (16.1)
Gender (%)			
Male	89 (31.0)	234 (25.2)	342 (29.6)
Female	193 (67.3)	688 (74.0)	806 (69.8)
Non-binary	5 (1.7)	5 (0.5)	5 (0.4)
Ethnicity (n, %)[5]			
Hispanic	44 (15.3)	130 (14.0)	245 (21.2)
Non-Hispanic	228 (79.4)	788 (85.1)	886 (76.7)
BMI [kg/m^2, mean (SD)]	28.0 (6.6)	27.9 (6.9)	28.5 (6.9)
BMI (kg/m^2)			
< 18.5	2 (0.7)	30 (3.2)	32 (2.8)
18.5 – 24.9	105 (36.6)	348 (37.8)	375 (32.5)
25.0 – 29.9	97 (33.8)	271 (29.5)	344 (29.8)
30.0 – 39.9	59 (20.6)	216 (23.5)	313 (27.1)
≥ 40	19 (6.6)	55 (6.0)	83 (7.2)
Smoking status (n, %)			
Never	260 (90.6)	869 (93.9)	1079 (93.4)
Occasionally	11 (3.8)	28 (3.0)	41 (3.6)
Regularly	11 (3.8)	29 (3.1)	25 (2.2)

[1]All participants in CoVHORT who did not have a COVID-19 test result; [2]PCR negative; [3]PCR-positive; [4] Ethnicity: Prefer not to answer (n=2), Missing (n=13); Smoking status: Missing (n=5); BMI: missing (n=5) [5]Gender: Non-binary gender includes any reported gender other than male or female, including transgender. Prefer not to answer (n=1), Transgender male (n=2); Ethnicity: Prefer not to answer (n=8), Missing (n=4); Smoking status: Missing (n=4);); BMI: missing (n=10) [6] Ethnicity: Prefer not to answer (n=12), Missing (n=12); Smoking status: Missing (n=10);); BMI: missing (n=8)

Table 2. Characteristics of COVID-19 positive study participants (n=1,355) by reported number of COVID-19 disease symptoms.

Characteristics at study entry	No symptoms[1] (n = 180)	Any symptoms[2] (n= 1,155)	1-6 symptoms (n= 486)	7-9 symptoms (n= 364)	10-18 symptoms (n= 305)	p-value[3]
Age [years, mean (SD)]	41.9 (17.5)	43.9 (16.1)	44.5 (16.7)	45.1 (16.1)	41.6 (14.6)	0.03
Days since symptoms began [mean (SD)][4]	-	86.6 (95.7)	87.0 (100.7)	84.7 (87.8)	88.2 (97.1)	0.95
Days since positive test [mean (SD)][5]	68.5 (73.4)	83.3 (91.4)	83.3 (94.9)	82.7 (86.0)	84.3 (92.2)	0.88
Sex (n, %)[6]						< 0.001
Female	109 (60.6)	806 (69.8)	310 (63.8)	249 (68.4)	247 (81.0)	
Male	68 (37.8)	342 (29.6)	174 (35.8)	113 (31.0)	55 (18.0)	
Ethnicity (n, %)						0.05
Non-Hispanic	129 (72.1)	886 (77.5)	380 (79.0)	284 (79.3)	222 (73.0)	
Hispanic	48 (26.8)	245 (21.4)	94 (19.5)	72 (20.1)	79 (26.0)	
BMI [kg/m^2, mean (SD)]	27.6 (6.5)	28.5 (6.9)	27.4 (5.9)	29.0 (7.2)	29.6 (7.8)	< 0.001
BMI (kg/m^2)						
< 18.5	7 (3.9)	32 (2.8)	14 (2.9)	11 (3.0)	7 (2.3)	0.95
18.5 – 24.9	71 (39.4)	375 (32.7)	175 (36.3)	107 (29.6)	93 (30.7)	ref
25.0 – 29.9	45 (25.0)	344 (30.0)	159 (33.0)	108 (29.8)	77 (25.4)	0.81
30.0 – 39.9	46 (25.6)	313 (27.3)	116 (24.1)	103 (28.5)	94 (31.0)	0.01
≥ 40	8 (4.4)	83 (7.2)	18 (3.8)	33 (9.1)	32 (10.6)	< 0.001
Smoking Status						0.38
Never	172 (95.6)	1079 (94.2)	457 (95.2)	338 (93.1)	284 (94.0)	
Occasionally or Regularly	8 (4.5)	66 (5.8)	23 (4.8)	25 (6.9)	18 (6.0)	
Self-rated Severity Score[7]	-		3.6 (2.3)	5.6 (2.1)	6.8 (1.9)	< 0.001

[1]Sex: Nonbinary (n=1), Transgender male (n=1); Missing values or prefer not to Answer: Days since positive test (n =2), Ethnicity (n= 3), BMI (n=3). [2]Sex: Nonbinary (n=5), Transgender male (n=1), Transgender female (n=1); Missing values or prefer not to Answer: Days since symptom began (n =178), Days since positive test (n =3), Ethnicity (n= 24), BMI (n=8), Smoking Status (n=10). [3]P-values calculated using ordered logistic regression. [4]Number of days between start of symptoms and survey completion. [5]Number of days between positive test date and survey completion.[6]Non-binary is a term for gender identities that fall outside of the traditional gender binary of male and female, and is how several participants self-identified.

Table 3. Symptom characteristics of CoVHORT participants by case status, adjusted by age, sex, ethnicity, BMI, and smoking status.

Reported symptoms at study entry	COVID-19 positive[1] n= 1,155 n (%)	Untested participants[2] n= 288 n (%)	COVID-19 negative[3] n= 930 n (%)	Positive vs Untested OR (95% C I)	Positive vs Negative OR (95% CI)
Fatigue	957 (82.9)	236 (82.2)	680 (73.1)	1.1 (0.7, 1.5)	1.81 (1.5, 2.3)
Headache	861 (74.6)	167 (58.2)	495 (53.2)	2.1 (1.5, 2.7)	2.6 (2.2, 3.2)
Aches and pains or sore muscles	766 (66.3)	178 (62.0)	506 (54.4)	1.2 (0.9, 1.6)	1.7 (1.4, 2.0)
Loss of smell/taste	725 (62.8)	67 (23.4)	119 (12.8)	6.0 (4.3, 8.3)	12.4 (9.8, 15.7)
Cough	716 (61.9)	209 (72.8)	540 (58.1)	0.7 (0.5, 0.9)	1.2 (1.0, 1.4)
Fever	610 (52.8)	171 (59.6)	452 (48.6)	0.8 (0.6, 1.0)	1.2 (1.0, 1.5)
Runny nose/cold-like symptoms	684 (59.2)	139 (48.4)	451 (48.5)	1.5 (1.1, 1.9)	1.6 (1.3, 1.9)
Chills	563 (48.7)	132 (46.0)	339 (36.5)	1.2 (0.9, 1.5)	1.7 (1.4, 2.0)
Sore throat	543 (47.0)	161 (56.1)	507 (54.5)	0.6 (0.5, 0.8)	0.7 (0.6, 0.9)
Difficulty breathing or shortness of breath	475 (41.1)	132 (46.0)	319 (34.3)	0.8 (0.6, 1.1)	1.4 (1.1, 1.6)
Diarrhea	348 (30.1)	71 (24.7)	162 (17.4)	1.3 (1.0, 1.8)	2.0 (1.6, 2.5)
Nausea	326 (28.2)	51 (17.8)	129 (13.9)	1.9 (1.3, 2.7)	2.5 (1.9, 3.1)
Chest pain or pressure	362 (31.3)	88 (30.7)	246 (26.5)	1.1 (0.8, 1.4)	1.3 (1.1, 1.6)
Bone pain/nerve pain	212 (18.4)	20 (7.0)	66 (7.1)	2.9 (1.8, 4.8)	3.0 (2.2, 4.0)
Vomiting	92 (8.0)	13 (4.5)	44 (4.7)	1.8 (1.0, 3.5)	1.7 (1.1, 2.4)
Other	101 (8.7)	17 (5.9)	32 (3.4)	1.8 (1.0, 3.1)	2.9 (1.9, 4.3)
Rash on skin	82 (7.1)	15 (5.2)	38 (4.1)	1.5 (0.8, 2.8)	1.8 (1.2, 2.8)
Discoloration of fingers/toes	29 (2.5)	4 (1.4)	18 (1.9)	2.2 (0.6, 7.2)	1.3 (0.7, 2.3)
Loss of speech or movement	12 (1.0)	1 (0.4)	7 (0.8)	2.8 (0.4, 22.0)	1.3 (0.5, 3.4)
Conjunctivitis	26 (2.3)	11 (3.8)	28 (3.0)	0.6 (0.3, 1.3)	0.8 (0.5, 1.4)

[1]PCR-positive cases; [2]participants in CoVHORT who do not have a laboratory-confirmed result; [3]PCR or antibody negative.

Ethics Statement: This study involving human participants was reviewed and approved by the Institutional Review Board of the University of Arizona Human Subjects Protection Program (#2003521636A00). Written informed consent to participate in this study was provided by the participants or the participants' legal guardian/next of kin.

Funding Statement: This work was supported by the BIO5 Institute at The University of Arizona. Grant Number: N/A.

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Data Sharing: No additional data available.

Competing Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential competing interest.

Contributorship statement: KP-B, LF, MJ, MB, ZC, YK, KE, EJ, PG-F, SRC, LK and FC-M conceptualized the study and developed the initial study protocol. SK, KH, CC, EA, JH, EJ, CK and KP-B participated in the design of the protocol, and the drafting and reviewing of the manuscript. All authors critically reviewed the draft of the manuscript and approved the final version. All authors meet the criteria for authorship as developed by the International Committee for Medical Journal Editors.

References

1. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. New England Journal of Medicine. 2020;382(8):727-33.
2. Guan WJ, Ni ZY, Hu Y, Liang WH, Ou CQ, He JX, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. N Engl J Med. 2020;382(18):1708-20.
3. Petrilli CM, Jones SA, Yang J, Rajagopalan H, O'Donnell L, Chernyak Y, et al. Factors associated with hospital admission and critical illness among 5279 people with coronavirus disease 2019 in New York City: prospective cohort study. BMJ. 2020;369:m1966.
4. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet. 2020;395(10229):1054-62.
5. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. Lancet. 2020;395(10223):507-13.
6. Cai H. Sex difference and smoking predisposition in patients with COVID-19. Lancet Respir Med. 2020;8(4):e20.
7. Hygiene NDoHaM. Age-adjusted rates of lab confirmed COVID-19 non-hospitalized cases, estimated non-fatal hospitalized cases, and persons known to have died per 100,000 by race/ethnicity group. Albany, NY2020.
8. Grant MC, Geoghegan L, Arbyn M, Mohammed Z, McGuinness L, Clarke EL, et al. The prevalence of symptoms in 24,410 adults infected by the novel coronavirus (SARS-CoV-2; COVID-19): A systematic review and meta-analysis of 148 studies from 9 countries. PloS one. 2020;15(6):e0234765.
9. Struyf T, Deeks JJ, Dinnes J, et al. Signs and symptoms to determine if a patient presenting in primary care or hospital outpatient settings has COVID-19. Cochrane Database Syst Rev. 2021;2021(2). doi:10.1002/14651858.CD013665.PUB2
10. Lechien JR, Chiesa-Estomba CM, Place S, Van Laethem Y, Cabaraux P, Mat Q, et al. Clinical and epidemiological characteristics of 1420 European patients with mild-to-moderate coronavirus disease 2019. J Intern Med. 2020;288(3):335-44.
11. Larsen JR, Martin MR, Martin JD, Kuhn P, Hicks JB. Modeling the Onset of Symptoms of COVID-19. Front Public Health. 2020;8:473.
12. Catalfamo CJ, Heslin KM, Shilen A, Khan SM, Hunsaker JR, Austhof E, et al. Design of the Arizona CoVHORT: A Population-Based COVID-19 Cohort. Frontiers in Public Health. 2021;9(21).
13. Dillman D, Smyth J, Christian L. Internet, Mail, and Mixed-Mode Surveys: The Tailored Design Method. Hoboken, NJ: Wiley & Sons; 2009.
14. Services ADoH. COVID-19 Data Dashboard Phoenix, Arizona: Arizona Department of Health Services; 2021 [Available from: https://www.azdhs.gov/preparedness/epidemiology-disease-control/infectious-disease-epidemiology/covid-19/dashboards/index.php.
15. Macias Gil R, Marcelin JR, Zuniga-Blanco B, Marquez C, Mathew T, Piggott DA. COVID-19 Pandemic: Disparate Health Impact on the Hispanic/Latinx Population in the United States. J Infect Dis. 2020;222(10):1592-5.
16. Smith JC, Sausville EL, Girish V, Yuan ML, Vasudevan A, John KM, et al. Cigarette Smoke Exposure and Inflammatory Signaling Increase the Expression of the SARS-CoV-2 Receptor ACE2 in the Respiratory Tract. Dev Cell. 2020;53(5):514-29 e3.
17. Burke RM, Killerby ME, Newton S, Ashworth CE, Berns AL, Brennan S, et al. Symptom Profiles of a Convenience Sample of Patients with COVID-19 - United States, January-April 2020. MMWR Morb Mortal Wkly Rep. 2020;69(28):904-8.

Supplemental Table 1: Sensitivity and Specificity of COVID-19 symptoms

Reported symptoms at study entry	COVID-19 positive[1] n= 1,155 n (%)	COVID-19 negative[3] n= 930 n (%)	Sensitivity	Specificity
Fatigue	957 (82.9)	680 (73.1)	82.9	26.9
Headache	861 (74.6)	495 (53.2)	74.6	46.8
Aches and pains or sore muscles	766 (66.3)	506 (54.4)	66.3	45.6
Loss of smell/taste	725 (62.8)	119 (12.8)	62.8	87.2
Cough	716 (61.9)	540 (58.1)	62.0	41.9
Fever	610 (52.8)	452 (48.6)	52.8	51.4
Runny nose/cold-like symptoms	684 (59.2)	451 (48.5)	59.2	51.5
Chills	563 (48.7)	339 (36.5)	48.7	63.6
Sore throat	543 (47.0)	507 (54.5)	47.0	45.5
Difficulty breathing or shortness of breath	475 (41.1)	319 (34.3)	41.1	65.7
Diarrhea	348 (30.1)	162 (17.4)	30.1	82.6
Nausea	326 (28.2)	129 (13.9)	28.2	86.1
Chest pain or pressure	362 (31.3)	246 (26.5)	31.3	73.6
Bone pain/nerve pain	212 (18.4)	66 (7.1)	18.4	92.9
Vomiting	92 (8.0)	44 (4.7)	8.0	95.3
Other	101 (8.7)	32 (3.4)	8.7	96.6
Rash on skin	82 (7.1)	38 (4.1)	7.1	95.9
Discoloration of fingers/toes	29 (2.5)	18 (1.9)	2.5	98.1
Loss of speech or movement	12 (1.0)	7 (0.8)	1.0	99.3
Conjunctivitis	26 (2.3)	28 (3.0)	2.3	97.0

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract (*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	1,3
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4,5
Objectives	3	State specific objectives, including any prespecified hypotheses	4,5
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up (*b*) For matched studies, give matching criteria and number of exposed and unexposed	5
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	5
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	5
Bias	9	Describe any efforts to address potential sources of bias	5,9
Study size	10	Explain how the study size was arrived at	5
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	5
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding (*b*) Describe any methods used to examine subgroups and interactions (*c*) Explain how missing data were addressed (*d*) If applicable, explain how loss to follow-up was addressed (*e*) Describe any sensitivity analyses	6
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed (b) Give reasons for non-participation at each stage (c) Consider use of a flow diagram	6
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders (b) Indicate number of participants with missing data for each variable of interest (c) Summarise follow-up time (eg, average and total amount)	6
Outcome data	15*	Report numbers of outcome events or summary measures over time	6.7

Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included (*b*) Report category boundaries when continuous variables were categorized (*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	6,7
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	6,7
Discussion			
Key results	18	Summarise key results with reference to study objectives	6,7
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	9
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	7-10
Generalisability	21	Discuss the generalisability (external validity) of the study results	7-10
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	10

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.

BMJ Open

Elucidating symptoms of COVID-19 illness in the Arizona CoVHORT, a longitudinal cohort study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-053403.R2
Article Type:	Original research
Date Submitted by the Author:	13-Dec-2021
Complete List of Authors:	Khan, Sana; The University of Arizona, Department of Epidemiology and Biostatistics Farland, Leslie V.; The University of Arizona, Department of Epidemiology and Biostatistics; The University of Arizona Cancer Center Catalfamo, Collin; The University of Arizona, Department of Epidemiology and Biostatistics Austhof, Erika; The University of Arizona, Department of Epidemiology and Biostatistics Bell, Melanie L.; The University of Arizona, Department of Epidemiology and Biostatistics Chen, Zhao; The University of Arizona, Department of Epidemiology and Biostatistics Cordova-Marks, Felina; The University of Arizona Cancer Center; The University of Arizona, Department of Health Promotion Sciences Ernst, Kacey; The University of Arizona, Department of Epidemiology and Biostatistics, College of Public Health Garcia-Filion, Pamela; The University of Arizona, Department of Biomedical Informatics Heslin, Kelly M.; The University of Arizona, Department of Epidemiology and Biostatistics Hoskinson, Joshua; The University of Arizona Cancer Center Jehn, Megan; Arizona State University Joseph, Emily C.S.; The University of Arizona Cancer Center Kelley, Connor; The University of Arizona, Department of Epidemiology and Biostatistics Klimentidis, Yann; The University of Arizona, Department of Epidemiology and Biostatistics Russo Carroll, Stephanie; The University of Arizona, Department of Community, Environment, and Policy; The University of Arizona, Native Nations Institute at the Udall Center for Studies in Public Policy Kohler, Lindsay; The University of Arizona, Department of Epidemiology and Biostatistics; The University of Arizona, Department of Health Promotion Sciences Pogreba-Brown, Kristen; The University of Arizona, Department of Epidemiology and Biostatistics Jacobs, Elizabeth; The University of Arizona, Department of Epidemiology and Biostatistics; The University of Arizona Cancer Center
Primary Subject Heading:	Infectious diseases

Secondary Subject Heading:	Epidemiology
Keywords:	COVID-19, INFECTIOUS DISEASES, EPIDEMIOLOGY

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Elucidating symptoms of COVID-19 illness in the Arizona CoVHORT, a longitudinal cohort study

Sana M. Khan[1]

Leslie V. Farland[1,2]

Collin J. Catalfamo[1]

Erika Austhof[1]

Melanie L. Bell[1]

Zhao Chen[1]

Felina Cordova-Marks[2,3]

Kacey C. Ernst[1]

Pamela Garcia-Filion[4]

Kelly M. Heslin[1]

Joshua Hoskinson[2]

Megan L Jehn[5]

Emily C.S. Joseph[2]

Connor P. Kelley[1]

Yann C. Klimentidis[1]

Stephanie Russo Carroll[6,7]

Lindsay N. Kohler[1,3]

Kristen Pogreba-Brown[1]

Elizabeth T. Jacobs[1,2]

[1]Department of Epidemiology and Biostatistics, Mel and Enid Zuckerman College of Public Health, The University of Arizona, Tucson, AZ, United States; [2]University of Arizona Cancer Center, Tucson, AZ, United States, [3]Department of Health Promotion Sciences, Mel and Enid Zuckerman College of Public Health, The University of Arizona, Tucson, AZ, United States; [4]Department of Biomedical Informatics, College of Medicine–Phoenix, The University of Arizona, Phoenix, AZ, United States, [5]School of Human Evolution and Social Change, Arizona State University, Tempe, AZ, United States; [6]Department of Community, Environment and Policy, Mel and Enid Zuckerman College of Public Health, The University of Arizona, Tucson, AZ, United States; [7] Native Nations Institute at the Udall Center for Studies in Public Policy, University of Arizona, Tucson, Arizona.

Corresponding Author

Sana M. Khan, MPH
Doctoral Student, Epidemiology
University of Arizona
Mel and Enid Zuckerman College of Public Health
sanakhan@email.arizona.edu

Keywords: SARS-CoV-2, COVID-19, symptoms, prospective cohort

Abstract

Objective: To elucidate the symptoms of laboratory-confirmed COVID-19 cases as compared to laboratory-confirmed negative individuals and to the untested general population among all participants who reported symptoms within a large, prospective cohort study. **Setting and Design:** This work was conducted within the framework of The Arizona CoVHORT, a longitudinal prospective cohort study conducted among Arizona residents. **Participants:** Eligible participants were any individual living in Arizona and were recruited from across Arizona via COVID-19 case investigations, participation in testing studies, and a postcard mailing effort. **Primary and Secondary Outcome Measures:** The primary outcome measure was a comparison of the type and frequency of symptoms between COVID-19 positive cases, tested but negative individuals, and the general untested population who reported experiencing symptoms consistent with COVID-19. **Results:** Of the 1,335 laboratory-confirmed COVID-19 cases, 180 (13.5%) reported having no symptoms. Of those that did report symptoms, the most commonly reported were fatigue (82.2%), headache (74.6%), aches, pains, or sore muscles (66.3%), loss of taste or smell (62.8,) and cough (61.9%). In adjusted logistic regression models, COVID-19 positive participants were more likely than negative participants to experience loss of taste and smell (OR 12.1; 95% CI 9.6-15.2); bone or nerve pain (OR 3.0; 95% CI 2.2 - 4.1), headache (OR: 2.6; 95% CI 2.2-3.2), nausea (OR: 2.4; 95% CI 1.9-3.1), or diarrhea (OR: 2.1; 95% CI 1.7-2.6). Fatigue (82.9) and headache (74.9) had the highest sensitivities among symptoms, while loss of taste or smell (87.2) and bone or nerve pain (92.9) had the high specificities among significant symptoms associated with COVID-19. **Conclusion:** When comparing confirmed COVID-19 cases with either confirmed negative or untested participants, the pattern of symptoms that discriminates SARS-CoV-2 infection from those arising from other potential circulating pathogens may differ from general reports of symptoms among cases alone.

Strengths and limitations of this study:

- To our knowledge, no prior research has compared the prevalence of non-specific symptoms such as headache, fever, and runny nose between confirmed COVID-19-positive cases, confirmed COVID-19-negative cases, and a general, untested comparison group.
- While we have a variety of recruitment methods, the majority of our population was recruited from COVID-19 case investigations, testing and vaccination centers; this may lead to a participant population with greater access to health services than the general population.
- We cannot know the COVID-19 status of the untested participants; it is possible that some had already been infected but were asymptomatic or exhibited few symptoms and were not captured using our study design.

Introduction

In late 2019, the novel coronavirus SARS-CoV-2 was first recognized in China among patients who presented with pneumonia and the first scientific report appeared shortly thereafter [1]. On March 11th, 2020, the World Health Organization declared COVID-19 a pandemic. The pathogen has had multiple impacts on individual and societal wellbeing arising from both biological effects of the virus and policy-based mitigation. The majority of those infected with acute COVID-19 will go on to recover, though approximately 10-20% of COVID-19 patients overall will develop a severe case of disease, and may suffer from stroke, pneumonia, or acute respiratory distress syndrome (ARDS) and require intensive care and ventilation [2, 3].

Individuals are likely be most infectious during the early phases of the disease, when symptoms may be comparatively mild; therefore, it is important to elucidate the reported symptom patterns of COVID-19 patients compared to both laboratory-confirmed negative individuals and population-based controls. Several risk factors have been associated with disease susceptibility and severity including increasing age [4], male sex [2, 5, 6], and current or former smoking [3], which may also affect symptomology. Further, important differences in disease incidence and severity by race and ethnicity have emerged, with Native Americans, African Americans, and Latinos having higher COVID-19 prevalence, hospitalization, and mortality rates compared to non-Hispanic whites [7]. It is presently not well known if reports of symptoms or symptom patterns vary by these factors as well.

A recent meta-analysis of over 24,000 patients across nine countries reported on COVID-19 symptom presentation. In this work, the most commonly reported symptoms among people with COVID-19 were fever (78% of COVID-19 patients reporting), cough (57%), and fatigue (31%) [8]. Additionally, a systematic review published in February 2021 aimed to assess the diagnostic accuracy of symptoms associated with COVID-19; this review identified 44 studies which in total included over 26,000 participants. The review found that among 84 symptoms, cough and fever had high sensitivities and could be used as a prompt for further COVID-19 testing. However, a limitation of the review article is potential selection bias due to their sample being selected from primarily clinical settings [9]. Additional work examining symptoms in an unselected population is necessary to determine the syndromic presentation of COVID-19 in the general population. Another study conducted among European patients (n=1420) with mild or moderate COVID-19 found that the most frequently reported symptoms were headache (70%), loss of smell (70%), and obstruction of the nasal passages (68%) [10]. The authors of a separate study, the objective of which was to develop a better symptom modeling algorithm to aid targeted testing, concluded that fever and cough should be used as the key symptoms for rapid COVID-19 screening given their high sensitivity [11]. However, a major limitation of studies conducted to date is the lack of comparison of patient-reported symptoms to those of uninfected individuals. To our knowledge, no prior research has compared the prevalence of non-specific symptoms such as headache, fever, and runny nose between confirmed COVID-19-positive cases, confirmed COVID-19-negative cases, and population-based comparison groups.

Since COVID-19 community transmission began, Arizona has experienced multiple, severe, COVID-19 surges, with more than 1.1 million infections and 21,000 COVID-19-related deaths as

of October 2021. To address this epidemiological challenge, in May 2020, we initiated a large, prospective cohort in Arizona of racially- and ethnically- diverse residents in order to rigorously investigate factors contributing to variability in natural COVID-19 disease history including incidence, progression, resolution, and chronic outcomes of infection [12]. This COVID-19 cohort, dubbed The Arizona CoVHORT, provides a rich data source for multiple areas of inquiry related to the pandemic. The objective of the present work was to determine which symptoms were reported with the greatest frequency among participants who tested positive for COVID-19 as compared to participants who tested negative for COVID-19 and untested participants, while controlling for potential confounders such as age, ethnicity, sex, BMI, and smoking status. The findings of this paper will aid in the identification of symptoms that differentiate COVID-19 from other circulating infections or conditions, such as allergies.

Materials and Methods

Study Participants

The overall goal of the CoVHORT is to continuously enroll Arizonans into a cohort study to track both the acute and long-term phases of infection with SARS-CoV-2. The present analysis includes data through October 1st, 2021. Several recruitment methods were employed, which have been described in detail previously [12]. Briefly, the primary sources of recruitment have been through case investigations in a partnership with the Arizona Department of Health Services and other research studies and testing sites at the University of Arizona and Arizona State University, both of which have allowed for inclusion of laboratory-confirmed COVID-19 positive and negative participants. By October 1st, 2021, a total of 493 COVID-19-positive participants had been recruited through health department case investigations and 901 through our partnerships with studies and testing sites in Arizona.

A comprehensive mailing list was purchased that provides information on 2.2 million residents in Arizona. To recruit the population-based comparison group, a total of 17,500 postcards were mailed to a simple random sample of Pima County, Arizona residents in July 2020. Consistent with the Dillman method to maximize participation and minimize bias [13], three phased mailings of recruitment postcards occurred every two weeks. Participant-provided information from returned surveys was used to exclude those who had already enrolled from subsequent phases of the mailing campaign. Each list was screened prior to each mailing to reduce the number of undeliverable postcards. We have completed all three phases of the mailing campaign in Pima County, with 17,294 postcards delivered in the first phase, 17,147 in the second phase, and another 17,081 in the third phase. Method of recruitment is recorded for all participants allowing sensitivity analyses to be conducted within subgroups.

Patient and Public Involvement

We encourage active participation from members of The Arizona CoVHORT. The public and members of the cohort are invited to webinars where they are able to provide input, ask questions, and speak with the projects' principal investigators. We regularly revisit our survey instruments to ensure they are reflecting feedback from participants and are centering their experiences and

priorities. Study findings are disseminated at our study website (covhort.arizona.edu), along with a regularly updated participant dashboard containing descriptive data of the cohort population.

Survey Instruments

All participants included in the CoVHORT were sent identical structured electronic questionnaires at upon study entry, regardless of COVID-19 status. All participants were first asked, "Since January 1, have you experienced a sudden illness that led you to believe you had COVID19?" If they answered "yes", all participants, were asked to indicate which symptoms they had experienced since January 2020 from a list based upon prior reports in the literature, as well as through an open-text field. Participants who respond "no" are not asked about symptomology and were not included in this analysis. Regardless of symptom status, all participants were then asked if they had tested for the virus that causes COVID-19 with a nasal swab, throat swab, or saliva test since January 2020. Participants were classified as untested, positive, or negative based on their results (Table 1). Information regarding health and medical history was collected, along with other demographic data, including age, sex, race, and ethnicity, as well as for weight, height, and smoking status. From these data, we calculated body mass index as (kg/m^2), and categorized participants as having a BMI of <25, ≥ 25-29.9, and ≥ 30, to aid in clinical interpretation, as well as reported BMI as a continuous variable (Table 2).

Statistical analysis

Data were analyzed to describe the COVID-19 symptoms, estimate the prevalence of individual symptoms, and identify differences among COVID-19-positive cases compared to COVID-19-negative individuals and untested participants. Individual variables were summarized and reported using appropriate statistical measures: mean [standard deviation (SD)] for continuous and percent (%) for categorical variables. Among those who tested positive for COVID-19, we compared the participant characteristics upon study entry and number of symptoms (0 symptoms, 1-6 symptoms, 7-9 symptoms, 10-16 symptoms) using ordered logistic regression and report p-values to explore factors associated with increasing severity. A logistic regression model was fit for each symptom to measure the association, as measured by odds ratios and 95% confidence intervals, with COVID-19-positive status after adjusting for age, sex, ethnicity, BMI, and smoking status. Confounders were selected based on background knowledge. Logistic models were performed using participants with complete data (n=1,370) for all variables in the model. Additionally, we included sensitivity and specificity estimates for each individual symptom (Supplemental Table 1). Statistical significance was defined as $p < 0.05$, with two-sided tests. Data analyses was conducted using Stata 16.0 (College Station, TX).

Results

As of October 1st, 2021, the Arizona CoVHORT study had enrolled a total of 7,012 participants, 2,373 (33.8%) of whom reported symptoms associated with COVID-19 since January 2021. Of these participants, 1,335 (56.3%) had lab-confirmed positive COVID-19 result, 930 (39.2%) had a lab-confirmed COVID-19 negative result, and 288 (12.1%) were untested (Table 1). The participants were majority female (70.4%) and white (89.4%) and had a mean (SD) age of 44.5 (15.3) years. COVID-19-positive participants were younger (43.9 years) than COVID-19-negative

participants (44.8 years), and participants who had not been tested for COVID-19 (46.8 years). COVID-19 positive participants were more likely to be Hispanic (22.2%), compared to COVID-19-negative participants (14.0%) and untested CoVHORT participants (16.0%). COVID-19-positive participants were more likely to have a body mass index (BMI) of greater than 30 kg/m^2 (34.3%) compared with COVID-19-negative participants (29.5%) and untested CoVHORT participants (27.2%). Of the 1,335 lab-confirmed COVID-19-positive participants, the majority (86.5%) reported having experienced at least one symptom after diagnosis, while the remaining 180 participants (13.5%) were asymptomatic, having reported never experiencing any symptoms (Table 2). When asked to self-rate the severity of their illness on a scale of 0-10, those who reported 10-18 symptoms reported a mean (SD) severity score of 6.8 (1.9), while participants with 7-9 symptoms reported a mean severity score of 5.6 (2.1), and participants with 1-6 symptoms reported a mean severity score of 3.6 (2.3) (Table 2). We assessed days since symptom onset and days since test date with the survey completion date and found no significant difference between symptom groups (Table 2).

As shown in Table 3, other common symptoms that lab-confirmed COVID-19-positive participants reported at any time in their disease course included fatigue (82.9%), headache (74.6%), loss of taste or smell (62.8%), aches and pains or sore muscles (66.3%), and cough (61.9%). COVID-19-positive participants had greater odds of reporting loss of taste or smell, bone or nerve pain, headache, nausea, and cold-like symptoms when compared to participants who tested negative for COVID-19 and participants who were never tested for COVID-19. While the magnitude of effect for these latter symptoms was smaller, all results were statistically significant. No differences between groups were observed for cough, fever, sore throat, loss of speech or movement, discoloration of fingers or toes, and conjunctivitis. After adjusting for age, ethnicity, sex, BMI, and smoking status, COVID-19 positive participants were more likely than negative participants to experience loss of taste and smell (OR 12.1, 95% CI 9.6-15.2), bone or nerve pain (OR 3.0, CI 2.2-4.1), headache (OR 2.6, CI 2.2-3.2), nausea (OR 2.4, CI 1.9-3.1), and diarrhea (OR 2.1, CI 1.7-2.6) (Table 3). Similarly, the symptoms with the strongest association when comparing COVID-19-positive cases with the untested participants were loss of taste or smell (OR 5.8, CI 4.2-7.9), bone/nerve pain (OR 2.9, CI 1.8-4.6), headache (OR 2.1, CI 1.6-2.7), nausea (OR 1.7, CI 1.2-2.5), and cold-like symptoms (OR 1.5, CI 1.1-2.0). Fatigue (82.9), headache (74.6), and aches and pains or sore muscles (66.3) were shown to have the highest sensitivities among symptoms, while loss of taste or smell (87.2) and bone or nerve pain (92.9) had high specificity among the significant symptoms (Supplemental Table 1).

Discussion

We assessed the type and frequency of symptoms between COVID-19 positive cases, tested but negative individuals, and the general untested population who reported experiencing symptoms consistent with COVID-19. We determined that lab-confirmed COVID-19 cases differed in age, ethnicity, BMI from COVID-negative participants, and untested cohort members. These same factors were associated with reported symptom severity. The most commonly reported first symptoms among COVID-19 positive participants were sore throat, followed by headache, cough, runny nose/cold-like symptoms, and fatigue. Discriminating symptoms for COVID-19-positivity

included loss of taste and smell and bone or nerve pain as demonstrated by specificity analyses; while fatigue, headache, and aches and pains or sore muscles were shown to have the highest sensitivities among symptoms.

Individuals identifying as Hispanic in CoVHORT constituted 33.5% of the recruited COVID-positive participants, mirroring the broader statewide case composition reported by the Arizona Department of Health Services [14]. By comparison, they constituted far fewer of the lab-negative and untested groups. As discussed by Macias Gil et al. [15], the burden of COVID-19 on communities of color has been far more extreme due to extant healthcare disparities, with greater rates of hospitalizations and deaths among U.S. Hispanics as compared to whites being reported in other studies [15]. Further, because publicly-available COVID-19 data by race or ethnicity may have missing values, it is critical to continue to follow up the health outcomes of this medically-vulnerable group.

Differences in disease outcomes by body size have been well-documented. In the first large study of COVID-19 patients in the United States, obesity was determined to be a major risk factor for hospitalization [3], but it remains unclear whether this finding is attributable to comorbidities that are themselves associated with both larger body size and with severe COVID-19. In the present work, only those with a BMI of greater than 30 kg/m^2 were at increased risk for being COVID-19 positive compared to those with classified as normal weight or overweight. Disentangling the drivers of susceptibility and disease progression will require long-term follow-up in a large, diverse study population, particularly as several comorbidities, such as type 2 diabetes, are also strongly associated with larger body size. Future work from this cohort will include detailed investigations of the impact of body size on susceptibility to and recovery from COVID-19.

Another equivocal risk factor is smoking, which to date has not been clearly demonstrated to convey an increased risk for severe disease [3]. In the present work, there was no difference in COVID-19 test status by smoking status. A previous study in the United States indicated that current or former smokers were less likely to be hospitalized with COVID-19, but that former smokers were more likely to go on to develop severe disease after hospitalization, and no differences in frequency of critical illness were observed for current smokers [3]. However, smoking is known to upregulate the production of the ACE2 receptor cells needed for SARS-CoV-2 to invade cells, though nicotine is known to block the ACE2 receptors [16]. This paradox complicates the relationship between smoking and COVID-19. and there is significant variability in the literature. Therefore, more work is needed to assess the role of smoking in COVID-19 disease progression, and future work from CoVHORT will include a detailed analysis of different smoking modalities such as vaping or e-cigarettes, cigar, and cigarette smoking.

Several efforts have been made to identify and characterize the symptoms associated with COVID-19 to allow for more efficient and targeted screening practices, as well as to differentiate SARS-CoV-2 infection from other diseases, such as influenza [8-10, 17]. However, these reports of COVID-19 symptoms have largely been confined to hospitalized or outpatient patient population and are lacking a symptomatic COVID-19 negative comparison group. Because many of the symptoms reported as being associated with COVID-19 are general symptoms that could be associated with conditions such as allergies or other infectious illnesses such as influenza, there is

an urgent need to evaluate the prevalence of reported symptoms of confirmed COVID-19-positive cases as compared to confirmed COVID-19-negative individuals, as well as with the prevalence of symptoms in the general population.

The results of the present study demonstrate that in southern Arizona, the most common first symptom reported by COVID-19-positive participants was sore throat, other common first symptoms of COVID-19 included headache, cough, runny nose or cold-like symptoms, and fatigue. While these are the same cluster of symptoms as reported by Larsen et al. in a large meta-analysis of more than 50,000 subjects, with data captured by the World Health Organization (WHO), the timing of appearance differed [11]. Specifically, the report by Larsen concluded that the order of symptom appearance was estimated to be fever, cough, nausea, and vomiting; while in the current work, the first symptom reported by the majority of cases was sore throat, followed by headache, cough, and runny nose; only 6% of participants had fever as their first symptom. Differences in the study population, including geographic location, sex, age, timing within the pandemic, severity of illness that prompted healthcare seeking behavior and testing, testing accessibility, and race differences across the spectrum of studies employed in the meta-analysis, may explain some of the inconsistent results for first reported symptoms.

An example of this variation in symptom reporting can be observed regarding the number of symptoms that women experienced as compared to men. Women were more likely to be classified in the category of the greatest number of symptoms than men, as were those with a BMI of greater than 30 kg/m^2, compared to those with a BMI below that threshold, although these findings were not statistically significant. A greater proportion of smokers was observed in the asymptomatic category, as compared to the any symptoms category. These findings suggest that ascertaining the type and order of COVID-19 specific symptomology may be confounded by characteristics of the participants.

With regard to overall COVID-19 symptoms, the greatest differences between laboratory-confirmed positive and negative participants were observed for loss of smell and taste and bone or nerve pain, followed by vomiting, nausea, and headache. A similar pattern was seen when comparing cases to the overall untested sample. To date, most work regarding symptoms has relied upon the frequency of symptom occurrence among cases, with little ability to ascertain the degree to which these symptoms differentiate cases from non-cases. For instance, the largest meta-analysis of COVID-19 symptomology to date included data from 24,410 cases from nine countries reported that the most common symptoms were fever (78%), cough (57%), and fatigue (31%) [8]. A smaller study within the United States found that the frequency of symptoms among cases was highest for cough (84%), fever (80%), aches and pains (63%), chills (63%), and fatigue (62%) [17]. In comparison, herein we found that the most common symptoms reported by cases were fatigue, headache, loss of smell or taste, cough, aches or pains, or sore muscles.

A key finding of this work is that the discrimination of COVID-19-positive symptoms from others requires comparison groups. General symptoms reported differ from those which may be applied to differentiate COVID-19 from other infectious diseases or conditions that are present in the underlying population. The symptoms that demonstrated the greatest difference between COVID-19-positive participants and the prevalence of symptoms among laboratory-confirmed COVID-19

negative participants or in the general population were loss of smell and taste, bone or nerve pain, headache, nausea, and fatigue.

The strengths of this study are its prospective nature, ability to capture data for laboratory-confirmed COVID-19-positive cases who have not been hospitalized, and the presence of comparison groups among both those who tested negative for COVID-19 as well as a population base drawn from throughout Arizona. These aspects allowed us to compare symptoms between cases and laboratory-confirmed uninfected individuals. However, limitations of the work must also be considered. First and foremost, while we are able to recruit participants via follow up to COVID-19 testing, participants' test results and symptoms are self-reported. Furthermore, although we have self-reported, laboratory-confirmed negative participants, we cannot know the COVID-19 status of the untested participants. It is possible that some had already been infected but were asymptomatic or exhibited few symptoms. This would likely attenuate any associations between exposure and outcomes in this study. It is also important to acknowledge that participants who indicated not experiencing symptoms that led them to believe they had COVID-19, regardless of testing status, were not asked to indicate which symptoms they had experienced. These participants were not able to directly indicate that they experienced no symptoms from the provided list; however, because the majority of CoVHORT participants who undergo testing for COVID-19 enroll at a time point after receiving their test results, we believe that the likelihood that participants who indicate not experiencing symptoms actually experienced symptoms from our list offered to participants who indicate experiencing symptoms is low. Additionally, there may be differences in the source population for cases as compared to the laboratory-negative participants and untested participants due to the differences in recruitment strategies for these populations. For example, while postcards were mailed to a random selection of households, it is possible Latinx participants were less likely to respond to this method than direct recruitment as cases during routine case follow-up. This could bias the association between being COVID-19-positive and Latinx away from the null. However, our race/ethnicity profile among cases is approximately similar to the overall distribution of cases throughout Arizona, suggesting a representative sample. Therefore, bias would potentially come from differential responses to other recruitment methods. This was an exploratory study, with a large number of statistical tests, and therefore care should be taken when considering p-values.

In conclusion, the findings of this analysis from the Arizona CoVHORT study show variation in several individual characteristics between COVID-19-positive participants, negative participants, and the untested population, which will be studied in future publications to assess the contributors to these observations. In addition, we found that in southern Arizona, COVID-19 positive participants most commonly reported a sore throat headache, fatigue, cough, or runny nose as the first symptom they noted. These results may aid in earlier identification of cases in the future and highlight the continued importance of addressing surveillance strategies as the pandemic continues.

Table 1. Demographic characteristics of CoVHORT participants who reported symptoms and were laboratory-confirmed positive for COVID-19, those who were tested and were negative for COVID-19, and those without COVID-19 test results in the CoVHORT population.

		Lab-confirmed COVID-19 status	
Characteristics at study entry	Untested participants[1,4] n= 288	COVID-19 negative[2,5] n= 930	COVID-19 positive[3,6] n= 1,335
Age [years, mean (SD)]	46.8 (14.2)	44.8 (14.0)	43.9 (16.1)
Gender (%)			
Male	89 (31.0)	234 (25.2)	342 (29.6)
Female	193 (67.3)	688 (74.0)	806 (69.8)
Non-binary	5 (1.7)	5 (0.5)	5 (0.4)
Ethnicity (n, %)[5]			
Hispanic	44 (15.3)	130 (14.0)	245 (21.2)
Non-Hispanic	228 (79.4)	788 (85.1)	886 (76.7)
BMI [kg/m^2, mean (SD)]	28.0 (6.6)	27.9 (6.9)	28.5 (6.9)
BMI (kg/m^2)			
< 18.5	2 (0.7)	30 (3.2)	32 (2.8)
18.5 – 24.9	105 (36.6)	348 (37.8)	375 (32.5)
25.0 – 29.9	97 (33.8)	271 (29.5)	344 (29.8)
30.0 – 39.9	59 (20.6)	216 (23.5)	313 (27.1)
≥ 40	19 (6.6)	55 (6.0)	83 (7.2)
Smoking status (n, %)			
Never	260 (90.6)	869 (93.9)	1079 (93.4)
Occasionally	11 (3.8)	28 (3.0)	41 (3.6)
Regularly	11 (3.8)	29 (3.1)	25 (2.2)

[1]All participants in CoVHORT who did not have a COVID-19 test result; [2]PCR negative; [3]PCR-positive; [4] Ethnicity: Prefer not to answer (n=2), Missing (n=13); Smoking status: Missing (n=5); BMI: missing (n=5) [5]Gender: Non-binary gender includes any reported gender other than male or female, including transgender. Prefer not to answer (n=1), Transgender male (n=2); Ethnicity: Prefer not to answer (n=8), Missing (n=4); Smoking status: Missing (n=4);); BMI: missing (n=10) [6] Ethnicity: Prefer not to answer (n=12), Missing (n=12); Smoking status: Missing (n=10);); BMI: missing (n=8)

Table 2. Characteristics of COVID-19 positive study participants (n=1,355) by reported number of COVID-19 disease symptoms.

Characteristics at study entry	No symptoms[1] (n = 180)	Any symptoms[2] (n= 1,155)	1-6 symptoms (n= 486)	7-9 symptoms (n= 364)	10-18 symptoms (n= 305)	p-value[3]
Age [years, mean (SD)]	41.9 (17.5)	43.9 (16.1)	44.5 (16.7)	45.1 (16.1)	41.6 (14.6)	0.03
Days since symptoms began [mean (SD)][4]	-	86.6 (95.7)	87.0 (100.7)	84.7 (87.8)	88.2 (97.1)	0.95
Days since positive test [mean (SD)][5]	68.5 (73.4)	83.3 (91.4)	83.3 (94.9)	82.7 (86.0)	84.3 (92.2)	0.88
Sex (n, %)[6]						< 0.001
Female	109 (60.6)	806 (69.8)	310 (63.8)	249 (68.4)	247 (81.0)	
Male	68 (37.8)	342 (29.6)	174 (35.8)	113 (31.0)	55 (18.0)	
Ethnicity (n, %)						0.05
Non-Hispanic	129 (72.1)	886 (77.5)	380 (79.0)	284 (79.3)	222 (73.0)	
Hispanic	48 (26.8)	245 (21.4)	94 (19.5)	72 (20.1)	79 (26.0)	
BMI [kg/m^2, mean (SD)]	27.6 (6.5)	28.5 (6.9)	27.4 (5.9)	29.0 (7.2)	29.6 (7.8)	< 0.001
BMI (kg/m^2)						
< 18.5	7 (3.9)	32 (2.8)	14 (2.9)	11 (3.0)	7 (2.3)	0.95
18.5 – 24.9	71 (39.4)	375 (32.7)	175 (36.3)	107 (29.6)	93 (30.7)	ref
25.0 – 29.9	45 (25.0)	344 (30.0)	159 (33.0)	108 (29.8)	77 (25.4)	0.81
30.0 – 39.9	46 (25.6)	313 (27.3)	116 (24.1)	103 (28.5)	94 (31.0)	0.01
≥ 40	8 (4.4)	83 (7.2)	18 (3.8)	33 (9.1)	32 (10.6)	< 0.001
Smoking Status						0.38
Never	172 (95.6)	1079 (94.2)	457 (95.2)	338 (93.1)	284 (94.0)	
Occasionally or Regularly	8 (4.5)	66 (5.8)	23 (4.8)	25 (6.9)	18 (6.0)	
Self-rated Severity Score[7]	-		3.6 (2.3)	5.6 (2.1)	6.8 (1.9)	< 0.001

[1]Sex: Nonbinary (n=1), Transgender male (n=1); Missing values or prefer not to Answer: Days since positive test (n =2), Ethnicity (n= 3), BMI (n=3). [2]Sex: Nonbinary (n=5), Transgender male (n=1), Transgender female (n=1); Missing values or prefer not to Answer: Days since symptom began (n =178), Days since positive test (n =3), Ethnicity (n= 24), BMI (n=8), Smoking Status (n=10). [3]P-values calculated using ordered logistic regression. [4]Number of days between start of symptoms and survey completion. [5]Number of days between positive test date and survey completion.[6]Non-binary is a term for gender identities that fall outside of the traditional gender binary of male and female, and is how several participants self-identified.

Table 3. Symptom characteristics and odds ratios of CoVHORT participants using a logistic regression model adjusted for case status, age, sex, ethnicity, BMI, and smoking status.

Reported symptoms at study entry	COVID-19 positive[1] n= 1,155 n (%)	Untested participants[2] n= 288 n (%)	COVID-19 negative[3] n= 930 n (%)	Positive vs Untested OR (95% C I)	Positive vs Negative OR (95% CI)
Fatigue	957 (82.9)	236 (82.2)	680 (73.1)	1.1 (0.7, 1.5)	1.81 (1.5, 2.3)
Headache	861 (74.6)	167 (58.2)	495 (53.2)	2.1 (1.5, 2.7)	2.6 (2.2, 3.2)
Aches and pains or sore muscles	766 (66.3)	178 (62.0)	506 (54.4)	1.2 (0.9, 1.6)	1.7 (1.4, 2.0)
Loss of smell/taste	725 (62.8)	67 (23.4)	119 (12.8)	6.0 (4.3, 8.3)	12.4 (9.8, 15.7)
Cough	716 (61.9)	209 (72.8)	540 (58.1)	0.7 (0.5, 0.9)	1.2 (1.0, 1.4)
Fever	610 (52.8)	171 (59.6)	452 (48.6)	0.8 (0.6, 1.0)	1.2 (1.0, 1.5)
Runny nose/cold-like symptoms	684 (59.2)	139 (48.4)	451 (48.5)	1.5 (1.1, 1.9)	1.6 (1.3, 1.9)
Chills	563 (48.7)	132 (46.0)	339 (36.5)	1.2 (0.9, 1.5)	1.7 (1.4, 2.0)
Sore throat	543 (47.0)	161 (56.1)	507 (54.5)	0.6 (0.5, 0.8)	0.7 (0.6, 0.9)
Difficulty breathing or shortness of breath	475 (41.1)	132 (46.0)	319 (34.3)	0.8 (0.6, 1.1)	1.4 (1.1, 1.6)
Diarrhea	348 (30.1)	71 (24.7)	162 (17.4)	1.3 (1.0, 1.8)	2.0 (1.6, 2.5)
Nausea	326 (28.2)	51 (17.8)	129 (13.9)	1.9 (1.3, 2.7)	2.5 (1.9, 3.1)
Chest pain or pressure	362 (31.3)	88 (30.7)	246 (26.5)	1.1 (0.8, 1.4)	1.3 (1.1, 1.6)
Bone pain/nerve pain	212 (18.4)	20 (7.0)	66 (7.1)	2.9 (1.8, 4.8)	3.0 (2.2, 4.0)
Vomiting	92 (8.0)	13 (4.5)	44 (4.7)	1.8 (1.0, 3.5)	1.7 (1.1, 2.4)
Other	101 (8.7)	17 (5.9)	32 (3.4)	1.8 (1.0, 3.1)	2.9 (1.9, 4.3)
Rash on skin	82 (7.1)	15 (5.2)	38 (4.1)	1.5 (0.8, 2.8)	1.8 (1.2, 2.8)
Discoloration of fingers/toes	29 (2.5)	4 (1.4)	18 (1.9)	2.2 (0.6, 7.2)	1.3 (0.7, 2.3)
Loss of speech or movement	12 (1.0)	1 (0.4)	7 (0.8)	2.8 (0.4, 22.0)	1.3 (0.5, 3.4)
Conjunctivitis	26 (2.3)	11 (3.8)	28 (3.0)	0.6 (0.3, 1.3)	0.8 (0.5, 1.4)

[1]PCR-positive cases; [2]participants in CoVHORT who do not have a laboratory-confirmed result; [3]PCR or antibody negative.

Ethics Statement: This study involving human participants was reviewed and approved by the Institutional Review Board of the University of Arizona Human Subjects Protection Program (#2003521636A00). Written informed consent to participate in this study was provided by the participants or the participants' legal guardian/next of kin.

Funding Statement: This work was supported by the BIO5 Institute at The University of Arizona. Grant Number: N/A.

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Data Sharing: No additional data available.

Competing Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential competing interest.

Contributorship statement: KP-B, LF, MJ, MB, ZC, YK, KE, EJ, PG-F, SRC, LK and FC-M conceptualized the study and developed the initial study protocol. SK, KH, CC, EA, JH, EJ, CK and KP-B participated in the design of the protocol, and the drafting and reviewing of the manuscript. All authors critically reviewed the draft of the manuscript and approved the final version. All authors meet the criteria for authorship as developed by the International Committee for Medical Journal Editors.

References

1. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. New England Journal of Medicine. 2020;382(8):727-33.
2. Guan WJ, Ni ZY, Hu Y, Liang WH, Ou CQ, He JX, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. N Engl J Med. 2020;382(18):1708-20.
3. Petrilli CM, Jones SA, Yang J, Rajagopalan H, O'Donnell L, Chernyak Y, et al. Factors associated with hospital admission and critical illness among 5279 people with coronavirus disease 2019 in New York City: prospective cohort study. BMJ. 2020;369:m1966.
4. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet. 2020;395(10229):1054-62.
5. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. Lancet. 2020;395(10223):507-13.
6. Cai H. Sex difference and smoking predisposition in patients with COVID-19. Lancet Respir Med. 2020;8(4):e20.
7. Hygiene NDoHaM. Age-adjusted rates of lab confirmed COVID-19 non-hospitalized cases, estimated non-fatal hospitalized cases, and persons known to have died per 100,000 by race/ethnicity group. Albany, NY2020.
8. Grant MC, Geoghegan L, Arbyn M, Mohammed Z, McGuinness L, Clarke EL, et al. The prevalence of symptoms in 24,410 adults infected by the novel coronavirus (SARS-CoV-2; COVID-19): A systematic review and meta-analysis of 148 studies from 9 countries. PloS one. 2020;15(6):e0234765.
9. Struyf T, Deeks JJ, Dinnes J, et al. Signs and symptoms to determine if a patient presenting in primary care or hospital outpatient settings has COVID-19. Cochrane Database Syst Rev. 2021;2021(2). doi:10.1002/14651858.CD013665.PUB2
10. Lechien JR, Chiesa-Estomba CM, Place S, Van Laethem Y, Cabaraux P, Mat Q, et al. Clinical and epidemiological characteristics of 1420 European patients with mild-to-moderate coronavirus disease 2019. J Intern Med. 2020;288(3):335-44.
11. Larsen JR, Martin MR, Martin JD, Kuhn P, Hicks JB. Modeling the Onset of Symptoms of COVID-19. Front Public Health. 2020;8:473.
12. Catalfamo CJ, Heslin KM, Shilen A, Khan SM, Hunsaker JR, Austhof E, et al. Design of the Arizona CoVHORT: A Population-Based COVID-19 Cohort. Frontiers in Public Health. 2021;9(21).
13. Dillman D, Smyth J, Christian L. Internet, Mail, and Mixed-Mode Surveys: The Tailored Design Method. Hoboken, NJ: Wiley & Sons; 2009.
14. Services ADoH. COVID-19 Data Dashboard Phoenix, Arizona: Arizona Department of Health Services; 2021 [Available from: https://www.azdhs.gov/preparedness/epidemiology-disease-control/infectious-disease-epidemiology/covid-19/dashboards/index.php.
15. Macias Gil R, Marcelin JR, Zuniga-Blanco B, Marquez C, Mathew T, Piggott DA. COVID-19 Pandemic: Disparate Health Impact on the Hispanic/Latinx Population in the United States. J Infect Dis. 2020;222(10):1592-5.
16. Smith JC, Sausville EL, Girish V, Yuan ML, Vasudevan A, John KM, et al. Cigarette Smoke Exposure and Inflammatory Signaling Increase the Expression of the SARS-CoV-2 Receptor ACE2 in the Respiratory Tract. Dev Cell. 2020;53(5):514-29 e3.
17. Burke RM, Killerby ME, Newton S, Ashworth CE, Berns AL, Brennan S, et al. Symptom Profiles of a Convenience Sample of Patients with COVID-19 - United States, January-April 2020. MMWR Morb Mortal Wkly Rep. 2020;69(28):904-8.

Supplemental Table 1: Sensitivity and Specificity of COVID-19 symptoms

Reported symptoms at study entry	COVID-19 positive[1] n= 1,155 n (%)	COVID-19 negative[3] n= 930 n (%)	Sensitivity	Specificity
Fatigue	957 (82.9)	680 (73.1)	82.9	26.9
Headache	861 (74.6)	495 (53.2)	74.6	46.8
Aches and pains or sore muscles	766 (66.3)	506 (54.4)	66.3	45.6
Loss of smell/taste	725 (62.8)	119 (12.8)	62.8	87.2
Cough	716 (61.9)	540 (58.1)	62.0	41.9
Fever	610 (52.8)	452 (48.6)	52.8	51.4
Runny nose/cold-like symptoms	684 (59.2)	451 (48.5)	59.2	51.5
Chills	563 (48.7)	339 (36.5)	48.7	63.6
Sore throat	543 (47.0)	507 (54.5)	47.0	45.5
Difficulty breathing or shortness of breath	475 (41.1)	319 (34.3)	41.1	65.7
Diarrhea	348 (30.1)	162 (17.4)	30.1	82.6
Nausea	326 (28.2)	129 (13.9)	28.2	86.1
Chest pain or pressure	362 (31.3)	246 (26.5)	31.3	73.6
Bone pain/nerve pain	212 (18.4)	66 (7.1)	18.4	92.9
Vomiting	92 (8.0)	44 (4.7)	8.0	95.3
Other	101 (8.7)	32 (3.4)	8.7	96.6
Rash on skin	82 (7.1)	38 (4.1)	7.1	95.9
Discoloration of fingers/toes	29 (2.5)	18 (1.9)	2.5	98.1
Loss of speech or movement	12 (1.0)	7 (0.8)	1.0	99.3
Conjunctivitis	26 (2.3)	28 (3.0)	2.3	97.0

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract (*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	1,3
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4,5
Objectives	3	State specific objectives, including any prespecified hypotheses	4,5
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up (*b*) For matched studies, give matching criteria and number of exposed and unexposed	5
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	5
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	5
Bias	9	Describe any efforts to address potential sources of bias	5,9
Study size	10	Explain how the study size was arrived at	5
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	5
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding (*b*) Describe any methods used to examine subgroups and interactions (*c*) Explain how missing data were addressed (*d*) If applicable, explain how loss to follow-up was addressed (*e*) Describe any sensitivity analyses	6
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed (b) Give reasons for non-participation at each stage (c) Consider use of a flow diagram	6
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders (b) Indicate number of participants with missing data for each variable of interest (c) Summarise follow-up time (eg, average and total amount)	6
Outcome data	15*	Report numbers of outcome events or summary measures over time	6.7

Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included (*b*) Report category boundaries when continuous variables were categorized (*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	6,7
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	6,7
Discussion			
Key results	18	Summarise key results with reference to study objectives	6,7
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	9
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	7-10
Generalisability	21	Discuss the generalisability (external validity) of the study results	7-10
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	10

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.